

# UTILITY PATENT APPLICATION TRANSMITTAL

Attorney Docket No.

210121.427C15

First Inventor or Application Identifier

Jiangchun Xu

Title

COMPOSITIONS AND METHODS FOR THE THERAPY  
AND DIAGNOSIS OF PROSTATE CANCER

Express Mail Label No.

EL615230015US

Only for nonprovisional applications under 37 CFR § 1.53(b)

## APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

## ADDRESS TO:

Box Patent Application  
Assistant Commissioner for Patents  
Washington, D.C. 202311. ☐ General Authorization Form & Fee Transmittal  
(Submit an original and a duplicate for fee processing)2. ☒ Specification [Total Pages] **201**  
(preferred arrangement set forth below)

- Descriptive Title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R & D
- Reference to Microfiche Appendix
- Background of the Invention

- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claim(s)
- Abstract of the Disclosure

3. ☒ Drawing(s) (35 USC 113) [Total Sheets] **16**4. Oath or Declaration [Total Pages] **1**

- a. ☐ Newly executed (original or copy)
- b. ☐ Copy from a prior application (37 CFR 1.63(d))  
(for continuation/divisional with Box 17 completed)
- i. ☐ DELETION OF INVENTOR(S)  
Signed statement attached deleting  
inventor(s) named in the prior application,  
see 37 CFR 1.63(d)(2) and 1.33(b)

5. ☐ Incorporation By Reference (useable if box 4b is  
checked) The entire disclosure of the prior application,  
from which a copy of the oath or declaration is supplied  
under Box 4b, is considered to be part of the disclosure of  
the accompanying application and is hereby incorporated  
by reference therein.6. ☐ Microfiche Computer Program (Appendix)7. Nucleotide and Amino Acid Sequence Submission  
(if applicable, all necessary)

- a. ☒ Computer-Readable Copy
- b. ☒ Paper Copy (identical to computer copy)
- c. ☒ Statement verifying identity of above copies

## ACCOMPANYING APPLICATION PARTS

8. ☐ Assignment Papers (cover sheet & document(s))9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney  
(when there is an assignee)10. ☐ English Translation Document (if applicable)11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS Citations12. ☐ Preliminary Amendment13. ☒ Return Receipt Postcard14. ☐ Small Entity Statement(s) ☐ Statement filed in prior application,  
Status still proper and desired15. ☐ Certified Copy of Priority Document(s)  
(if foreign priority is claimed)16. ☒ Other: Certificate of Express Mail

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information below and in a preliminary amendment

☐ Continuation ☐ Divisional ☒ Continuation-In-Part (CIP) of prior Application No.: not assignedPrior application information: Examiner not assigned Group / Art Unit not assigned☐ Claims the benefit of Provisional Application No. \_\_\_\_\_

## CORRESPONDENCE ADDRESS

Jane E. R. Potter  
Seed Intellectual Property Law Group PLLC  
701 Fifth Avenue, Suite 6300  
Seattle, Washington 98104-7092  
Phone: (206) 622-4900 Fax: (206) 682-6031

Respectfully submitted,

TYPED or PRINTED NAME Jane E. R. PotterSIGNATURE Jane E. R. PotterREGISTRATION NO. 33,332Date June 13, 2000



PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Jiangchun Xu, Bellevue, WA; Davin C. Dillon, Redmond, WA; Jennifer L. Mitcham, Redmond, WA; Susan L. Harlocker, Seattle, WA; Yuqiu Jiang, Kent, WA; Michael D. Kalos, Seattle, WA; Gary R. Fanger, Mill Creek, WA; Marc W. Retter, Carnation, WA; John A. Stolk, Bothell, WA; Craig H. Day, Seattle, WA; Thomas S. Vedvick, Bainbridge Island, WA; Darrick Carter, Seattle, WA; Samuel Li, Redmond, WA; Aijun Wang, Issaquah, WA; Yasir A. W. Skeiky, Seattle, WA; William Hepler, Seattle, WA; Robert A. Henderson, Seattle, WA

Filed : June 13, 2000

For : COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF PROSTATE CANCER

Docket No. : 210121.427C15

Date : June 13, 2000

Box Patent Application  
Assistant Commissioner for Patents  
Washington, DC 20231

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Assistant Commissioner for Patents:

I hereby certify that the enclosures listed below are being deposited with the United States Postal Service "EXPRESS MAIL Post Office to Addressee" service under 37 C.F.R. § 1.10, Mailing Label Certificate No. EL615230015US, on June 13, 2000, addressed to Box Patent Application, Assistant Commissioner for Patents, Washington, DC 20231.

Respectfully submitted,

Seed Intellectual Property Law Group PLLC

  
Amber Straub/Jeanette West/Susan Johnson

Enclosures:

Postcard  
Form PTO/SB/05  
Specification, Claims, Abstract (201 pages)  
16 Sheets of Drawings (Figures 1-12)  
Sequence Listing (357 pages)  
Declaration for Sequence Listing  
Diskette for Sequence Listing

u:\sharons\210121\427



COMPOSITIONS AND METHODS FOR THE THERAPY  
AND DIAGNOSIS OF PROSTATE CANCER

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Patent Application No. 5 09/\_\_\_\_\_, filed May 12, 2000, which is a continuation-in-part of U.S. Patent Application No. 09/568,100, filed May 9, 2000, which is a continuation-in-part of U.S. Patent Application No. 09/536,857, filed March 27, 2000, which is a continuation-in-part of U. S. Patent Application No. 09/483,672, filed January 14, 2000, which is a continuation-in-part of U.S. Patent Application No. 09/439,313, filed November 12, 1999, which is a 10 continuation-in-part of U.S. Patent Application No. 09/352,616, filed July 13, 1999, which is a continuation-in-part of U.S. Patent Application No. 09/288,946, filed April 9, 1999, which is a continuation-in-part of U.S. Patent Application No. 09/232,149, filed January 15, 1999, which is a continuation-in-part of U.S. Patent Application No. 09/159,812, filed September 23, 1998, which is a continuation-in-part of U.S. Patent Application No. 15 09/115,453, filed July 14, 1998, which is a continuation-in-part of U.S. Patent Application No. 09/030,607, filed February 25, 1998, which is a continuation-in-part of U.S. Patent Application No. 09/020,956, filed February 9, 1998, which is a continuation-in-part of U.S. Patent Application No. 08/904,804, filed August 1, 1997, which is a continuation-in-part of U.S. Patent Application No. 08/806,099, filed February 25, 1997.

20 TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer, such as prostate cancer. The invention is more specifically related to polypeptides comprising at least a portion of a prostate-specific protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in compositions 25 for prevention and treatment of prostate cancer, and for the diagnosis and monitoring of such cancers.



## BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method for prevention or treatment is currently available. Current therapies, which are generally based on a combination of chemotherapy or surgery and radiation, continue to prove inadequate in many patients.

Prostate cancer is the most common form of cancer among males, with an estimated incidence of 30% in men over the age of 50. Overwhelming clinical evidence shows that human prostate cancer has the propensity to metastasize to bone, and the disease appears to progress inevitably from androgen dependent to androgen refractory status, leading to increased patient mortality. This prevalent disease is currently the second leading cause of cancer death among men in the U.S.

In spite of considerable research into therapies for the disease, prostate cancer remains difficult to treat. Commonly, treatment is based on surgery and/or radiation therapy, but these methods are ineffective in a significant percentage of cases. Two previously identified prostate specific proteins - prostate specific antigen (PSA) and prostatic acid phosphatase (PAP) - have limited therapeutic and diagnostic potential. For example, PSA levels do not always correlate well with the presence of prostate cancer, being positive in a percentage of non-prostate cancer cases, including benign prostatic hyperplasia (BPH). Furthermore, PSA measurements correlate with prostate volume, and do not indicate the level of metastasis.

In spite of considerable research into therapies for these and other cancers, prostate cancer remains difficult to diagnose and treat effectively. Accordingly, there is a need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.



## SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as prostate cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a prostate-specific protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789; (b) variants of a sequence recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789; and (c) complements of a sequence of (a) or (b). In specific embodiments, the polypeptides of the present invention comprise at least a portion of a tumor protein that includes an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778 and 780, and variants thereof.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a prostate-specific protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, immunogenic compositions, or vaccines for prophylactic or therapeutic use are provided. Such



compositions comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a prostate-specific protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, immunogenic compositions, or vaccines, are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Compositions are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a composition as recited above. The patient may be afflicted with prostate cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a prostate-specific protein, wherein the step of



contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as  
5 described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a prostate-specific protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide;  
10 under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective  
15 amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a prostate-specific protein; (ii) a polynucleotide encoding  
20 such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for  
25 determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.



Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be prostate cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of:

5 (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step

10 (b) and therefrom monitoring the progression of the cancer in the patient.

The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein; (b) detecting in the

15 sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one

20 oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

25 In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample



obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as  
 5 monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon  
 reference to the following detailed description and attached drawings. All references  
 10 disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

#### BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE IDENTIFIERS

Figure 1 illustrates the ability of T cells to kill fibroblasts expressing the  
 representative prostate-specific polypeptide P502S, as compared to control fibroblasts. The  
 15 percentage lysis is shown as a series of effector:target ratios, as indicated.

Figures 2A and 2B illustrate the ability of T cells to recognize cells  
 expressing the representative prostate-specific polypeptide P502S. In each case, the  
 number of  $\gamma$ -interferon spots is shown for different numbers of responders. In Figure 2A,  
 data is presented for fibroblasts pulsed with the P2S-12 peptide, as compared to fibroblasts  
 20 pulsed with a control E75 peptide. In Figure 2B, data is presented for fibroblasts  
 expressing P502S, as compared to fibroblasts expressing HER-2/*neu*.

Figure 3 represents a peptide competition binding assay showing that the  
 P1S#10 peptide, derived from P501S, binds HLA-A2. Peptide P1S#10 inhibits HLA-A2  
 restricted presentation of fluM58 peptide to CTL clone D150M58 in TNF release bioassay.  
 25 D150M58 CTL is specific for the HLA-A2 binding influenza matrix peptide fluM58.

Figure 4 illustrates the ability of T cell lines generated from P1S#10  
 immunized mice to specifically lyse P1S#10-pulsed Jurkat A2Kb targets and P501S-



transduced Jurkat A2Kb targets, as compared to EGFP-transduced Jurkat A2Kb. The percent lysis is shown as a series of effector to target ratios, as indicated.

Figure 5 illustrates the ability of a T cell clone to recognize and specifically lyse Jurkat A2Kb cells expressing the representative prostate-specific polypeptide P501S, thereby demonstrating that the P1S#10 peptide may be a naturally processed epitope of the P501S polypeptide.

Figures 6A and 6B are graphs illustrating the specificity of a CD8<sup>+</sup> cell line (3A-1) for a representative prostate-specific antigen (P501S). Figure 6A shows the results of a <sup>51</sup>Cr release assay. The percent specific lysis is shown as a series of effector:target ratios, as indicated. Figure 6B shows the production of interferon-gamma by 3A-1 cells stimulated with autologous B-LCL transduced with P501S, at varying effector:target ratios as indicated.

Figure 7 is a Western blot showing the expression of P501S in baculovirus.

Figure 8 illustrates the results of epitope mapping studies on P501S.

Figure 9 is a schematic representation of the P501S protein showing the location of transmembrane domains and predicted intracellular and extracellular domains.

Figure 10 is a genomic map showing the location of the prostate genes P775P, P704P, B305D, P712P and P774P within the Cat Eye Syndrome region of chromosome 22q11.2

Figure 11 shows the results of an ELISA assay to determine the specificity of rabbit polyclonal antisera raised against P501S.

Figures 12A(1), 12A(2), 12A(3), and B are the full-length cDNA (SEQ ID NO:591) and predicted amino acid (SEQ ID NO:592) sequences, respectively, for the clone P788P.

SEQ ID NO: 1 is the determined cDNA sequence for F1-13  
 SEQ ID NO: 2 is the determined 3' cDNA sequence for F1-12  
 SEQ ID NO: 3 is the determined 5' cDNA sequence for F1-12  
 SEQ ID NO: 4 is the determined 3' cDNA sequence for F1-16  
 SEQ ID NO: 5 is the determined 3' cDNA sequence for H1-1



	SEQ ID NO: 6 is the determined 3' cDNA sequence for H1-9
	SEQ ID NO: 7 is the determined 3' cDNA sequence for H1-4
	SEQ ID NO: 8 is the determined 3' cDNA sequence for J1-17
	SEQ ID NO: 9 is the determined 5' cDNA sequence for J1-17
5	SEQ ID NO: 10 is the determined 3' cDNA sequence for L1-12
	SEQ ID NO: 11 is the determined 5' cDNA sequence for L1-12
	SEQ ID NO: 12 is the determined 3' cDNA sequence for N1-1862
	SEQ ID NO: 13 is the determined 5' cDNA sequence for N1-1862
	SEQ ID NO: 14 is the determined 3' cDNA sequence for J1-13
10	SEQ ID NO: 15 is the determined 5' cDNA sequence for J1-13
	SEQ ID NO: 16 is the determined 3' cDNA sequence for J1-19
	SEQ ID NO: 17 is the determined 5' cDNA sequence for J1-19
	SEQ ID NO: 18 is the determined 3' cDNA sequence for J1-25
	SEQ ID NO: 19 is the determined 5' cDNA sequence for J1-25
15	SEQ ID NO: 20 is the determined 5' cDNA sequence for J1-24
	SEQ ID NO: 21 is the determined 3' cDNA sequence for J1-24
	SEQ ID NO: 22 is the determined 5' cDNA sequence for K1-58
	SEQ ID NO: 23 is the determined 3' cDNA sequence for K1-58
	SEQ ID NO: 24 is the determined 5' cDNA sequence for K1-63
20	SEQ ID NO: 25 is the determined 3' cDNA sequence for K1-63
	SEQ ID NO: 26 is the determined 5' cDNA sequence for L1-4
	SEQ ID NO: 27 is the determined 3' cDNA sequence for L1-4
	SEQ ID NO: 28 is the determined 5' cDNA sequence for L1-14
	SEQ ID NO: 29 is the determined 3' cDNA sequence for L1-14
25	SEQ ID NO: 30 is the determined 3' cDNA sequence for J1-12
	SEQ ID NO: 31 is the determined 3' cDNA sequence for J1-16
	SEQ ID NO: 32 is the determined 3' cDNA sequence for J1-21
	SEQ ID NO: 33 is the determined 3' cDNA sequence for K1-48
	SEQ ID NO: 34 is the determined 3' cDNA sequence for K1-55



SEQ ID NO: 35 is the determined 3' cDNA sequence for L1-2  
 SEQ ID NO: 36 is the determined 3' cDNA sequence for L1-6  
 SEQ ID NO: 37 is the determined 3' cDNA sequence for N1-1858  
 SEQ ID NO: 38 is the determined 3' cDNA sequence for N1-1860  
 5 SEQ ID NO: 39 is the determined 3' cDNA sequence for N1-1861  
 SEQ ID NO: 40 is the determined 3' cDNA sequence for N1-1864  
 SEQ ID NO: 41 is the determined cDNA sequence for P5  
 SEQ ID NO: 42 is the determined cDNA sequence for P8  
 SEQ ID NO: 43 is the determined cDNA sequence for P9  
 10 SEQ ID NO: 44 is the determined cDNA sequence for P18  
 SEQ ID NO: 45 is the determined cDNA sequence for P20  
 SEQ ID NO: 46 is the determined cDNA sequence for P29  
 SEQ ID NO: 47 is the determined cDNA sequence for P30  
 SEQ ID NO: 48 is the determined cDNA sequence for P34  
 15 SEQ ID NO: 49 is the determined cDNA sequence for P36  
 SEQ ID NO: 50 is the determined cDNA sequence for P38  
 SEQ ID NO: 51 is the determined cDNA sequence for P39  
 SEQ ID NO: 52 is the determined cDNA sequence for P42  
 SEQ ID NO: 53 is the determined cDNA sequence for P47  
 20 SEQ ID NO: 54 is the determined cDNA sequence for P49  
 SEQ ID NO: 55 is the determined cDNA sequence for P50  
 SEQ ID NO: 56 is the determined cDNA sequence for P53  
 SEQ ID NO: 57 is the determined cDNA sequence for P55  
 SEQ ID NO: 58 is the determined cDNA sequence for P60  
 25 SEQ ID NO: 59 is the determined cDNA sequence for P64  
 SEQ ID NO: 60 is the determined cDNA sequence for P65  
 SEQ ID NO: 61 is the determined cDNA sequence for P73  
 SEQ ID NO: 62 is the determined cDNA sequence for P75  
 SEQ ID NO: 63 is the determined cDNA sequence for P76



SEQ ID NO: 64 is the determined cDNA sequence for P79

SEQ ID NO: 65 is the determined cDNA sequence for P84

SEQ ID NO: 66 is the determined cDNA sequence for P68

SEQ ID NO: 67 is the determined cDNA sequence for P80 (also referred to

5 as P704P)

SEQ ID NO: 68 is the determined cDNA sequence for P82

SEQ ID NO: 69 is the determined cDNA sequence for U1-3064

SEQ ID NO: 70 is the determined cDNA sequence for U1-3065

SEQ ID NO: 71 is the determined cDNA sequence for V1-3692

10 SEQ ID NO: 72 is the determined cDNA sequence for 1A-3905

SEQ ID NO: 73 is the determined cDNA sequence for V1-3686

SEQ ID NO: 74 is the determined cDNA sequence for R1-2330

SEQ ID NO: 75 is the determined cDNA sequence for 1B-3976

SEQ ID NO: 76 is the determined cDNA sequence for V1-3679

15 SEQ ID NO: 77 is the determined cDNA sequence for 1G-4736

SEQ ID NO: 78 is the determined cDNA sequence for 1G-4738

SEQ ID NO: 79 is the determined cDNA sequence for 1G-4741

SEQ ID NO: 80 is the determined cDNA sequence for 1G-4744

SEQ ID NO: 81 is the determined cDNA sequence for 1G-4734

20 SEQ ID NO: 82 is the determined cDNA sequence for 1H-4774

SEQ ID NO: 83 is the determined cDNA sequence for 1H-4781

SEQ ID NO: 84 is the determined cDNA sequence for 1H-4785

SEQ ID NO: 85 is the determined cDNA sequence for 1H-4787

SEQ ID NO: 86 is the determined cDNA sequence for 1H-4796

25 SEQ ID NO: 87 is the determined cDNA sequence for 1I-4807

SEQ ID NO: 88 is the determined cDNA sequence for 1I-4810

SEQ ID NO: 89 is the determined cDNA sequence for 1I-4811

SEQ ID NO: 90 is the determined cDNA sequence for 1J-4876

SEQ ID NO: 91 is the determined cDNA sequence for 1K-4884



5

10

15

20

25



SEQ ID NO: 116 is the determined cDNA sequence for P90  
 SEQ ID NO: 117 is the determined cDNA sequence for P92  
 SEQ ID NO: 118 is the determined cDNA sequence for P95  
 SEQ ID NO: 119 is the determined cDNA sequence for P98  
 5 SEQ ID NO: 120 is the determined cDNA sequence for P102  
 SEQ ID NO: 121 is the determined cDNA sequence for P110  
 SEQ ID NO: 122 is the determined cDNA sequence for P111  
 SEQ ID NO: 123 is the determined cDNA sequence for P114  
 SEQ ID NO: 124 is the determined cDNA sequence for P115  
 10 SEQ ID NO: 125 is the determined cDNA sequence for P116  
 SEQ ID NO: 126 is the determined cDNA sequence for P124  
 SEQ ID NO: 127 is the determined cDNA sequence for P126  
 SEQ ID NO: 128 is the determined cDNA sequence for P130  
 SEQ ID NO: 129 is the determined cDNA sequence for P133  
 15 SEQ ID NO: 130 is the determined cDNA sequence for P138  
 SEQ ID NO: 131 is the determined cDNA sequence for P143  
 SEQ ID NO: 132 is the determined cDNA sequence for P151  
 SEQ ID NO: 133 is the determined cDNA sequence for P156  
 SEQ ID NO: 134 is the determined cDNA sequence for P157  
 20 SEQ ID NO: 135 is the determined cDNA sequence for P166  
 SEQ ID NO: 136 is the determined cDNA sequence for P176  
 SEQ ID NO: 137 is the determined cDNA sequence for P178  
 SEQ ID NO: 138 is the determined cDNA sequence for P179  
 SEQ ID NO: 139 is the determined cDNA sequence for P185  
 25 SEQ ID NO: 140 is the determined cDNA sequence for P192  
 SEQ ID NO: 141 is the determined cDNA sequence for P201  
 SEQ ID NO: 142 is the determined cDNA sequence for P204  
 SEQ ID NO: 143 is the determined cDNA sequence for P208  
 SEQ ID NO: 144 is the determined cDNA sequence for P211



SEQ ID NO: 145 is the determined cDNA sequence for P213  
 SEQ ID NO: 146 is the determined cDNA sequence for P219  
 SEQ ID NO: 147 is the determined cDNA sequence for P237  
 SEQ ID NO: 148 is the determined cDNA sequence for P239  
 5 SEQ ID NO: 149 is the determined cDNA sequence for P248  
 SEQ ID NO: 150 is the determined cDNA sequence for P251  
 SEQ ID NO: 151 is the determined cDNA sequence for P255  
 SEQ ID NO: 152 is the determined cDNA sequence for P256  
 SEQ ID NO: 153 is the determined cDNA sequence for P259  
 10 SEQ ID NO: 154 is the determined cDNA sequence for P260  
 SEQ ID NO: 155 is the determined cDNA sequence for P263  
 SEQ ID NO: 156 is the determined cDNA sequence for P264  
 SEQ ID NO: 157 is the determined cDNA sequence for P266  
 SEQ ID NO: 158 is the determined cDNA sequence for P270  
 15 SEQ ID NO: 159 is the determined cDNA sequence for P272  
 SEQ ID NO: 160 is the determined cDNA sequence for P278  
 SEQ ID NO: 161 is the determined cDNA sequence for P105  
 SEQ ID NO: 162 is the determined cDNA sequence for P107  
 SEQ ID NO: 163 is the determined cDNA sequence for P137  
 20 SEQ ID NO: 164 is the determined cDNA sequence for P194  
 SEQ ID NO: 165 is the determined cDNA sequence for P195  
 SEQ ID NO: 166 is the determined cDNA sequence for P196  
 SEQ ID NO: 167 is the determined cDNA sequence for P220  
 SEQ ID NO: 168 is the determined cDNA sequence for P234  
 25 SEQ ID NO: 169 is the determined cDNA sequence for P235  
 SEQ ID NO: 170 is the determined cDNA sequence for P243  
 SEQ ID NO: 171 is the determined cDNA sequence for P703P-DE1  
 SEQ ID NO: 172 is the predicted amino acid sequence for P703P-DE1  
 SEQ ID NO: 173 is the determined cDNA sequence for P703P-DE2



SEQ ID NO: 174 is the determined cDNA sequence for P703P-DE6  
 SEQ ID NO: 175 is the determined cDNA sequence for P703P-DE13  
 SEQ ID NO: 176 is the predicted amino acid sequence for P703P-DE13  
 SEQ ID NO: 177 is the determined cDNA sequence for P703P-DE14  
 5 SEQ ID NO: 178 is the predicted amino acid sequence for P703P-DE14  
 SEQ ID NO: 179 is the determined extended cDNA sequence for 1G-4736  
 SEQ ID NO: 180 is the determined extended cDNA sequence for 1G-4738  
 SEQ ID NO: 181 is the determined extended cDNA sequence for 1G-4741  
 SEQ ID NO: 182 is the determined extended cDNA sequence for 1G-4744  
 10 SEQ ID NO: 183 is the determined extended cDNA sequence for 1H-4774  
 SEQ ID NO: 184 is the determined extended cDNA sequence for 1H-4781  
 SEQ ID NO: 185 is the determined extended cDNA sequence for 1H-4785  
 SEQ ID NO: 186 is the determined extended cDNA sequence for 1H-4787  
 SEQ ID NO: 187 is the determined extended cDNA sequence for 1H-4796  
 15 SEQ ID NO: 188 is the determined extended cDNA sequence for 1I-4807  
 SEQ ID NO: 189 is the determined 3' cDNA sequence for 1I-4810  
 SEQ ID NO: 190 is the determined 3' cDNA sequence for 1I-4811  
 SEQ ID NO: 191 is the determined extended cDNA sequence for 1J-4876  
 SEQ ID NO: 192 is the determined extended cDNA sequence for 1K-4884  
 20 SEQ ID NO: 193 is the determined extended cDNA sequence for 1K-4896  
 SEQ ID NO: 194 is the determined extended cDNA sequence for 1G-4761  
 SEQ ID NO: 195 is the determined extended cDNA sequence for 1G-4762  
 SEQ ID NO: 196 is the determined extended cDNA sequence for 1H-4766  
 SEQ ID NO: 197 is the determined 3' cDNA sequence for 1H-4770  
 25 SEQ ID NO: 198 is the determined 3' cDNA sequence for 1H-4771  
 SEQ ID NO: 199 is the determined extended cDNA sequence for 1H-4772  
 SEQ ID NO: 200 is the determined extended cDNA sequence for 1D-4309  
 SEQ ID NO: 201 is the determined extended cDNA sequence for 1D.1-4278  
 SEQ ID NO: 202 is the determined extended cDNA sequence for 1D-4288



SEQ ID NO: 203 is the determined extended cDNA sequence for 1D-4283  
 SEQ ID NO: 204 is the determined extended cDNA sequence for 1D-4304  
 SEQ ID NO: 205 is the determined extended cDNA sequence for 1D-4296  
 SEQ ID NO: 206 is the determined extended cDNA sequence for 1D-4280  
 5 SEQ ID NO: 207 is the determined cDNA sequence for 10-d8fwd  
 SEQ ID NO: 208 is the determined cDNA sequence for 10-H10con  
 SEQ ID NO: 209 is the determined cDNA sequence for 11-C8rev  
 SEQ ID NO: 210 is the determined cDNA sequence for 7.g6fwd  
 SEQ ID NO: 211 is the determined cDNA sequence for 7.g6rev  
 10 SEQ ID NO: 212 is the determined cDNA sequence for 8-b5fwd  
 SEQ ID NO: 213 is the determined cDNA sequence for 8-b5rev  
 SEQ ID NO: 214 is the determined cDNA sequence for 8-b6fwd  
 SEQ ID NO: 215 is the determined cDNA sequence for 8-b6 rev  
 SEQ ID NO: 216 is the determined cDNA sequence for 8-d4fwd  
 15 SEQ ID NO: 217 is the determined cDNA sequence for 8-d9rev  
 SEQ ID NO: 218 is the determined cDNA sequence for 8-g3fwd  
 SEQ ID NO: 219 is the determined cDNA sequence for 8-g3rev  
 SEQ ID NO: 220 is the determined cDNA sequence for 8-h11rev  
 SEQ ID NO: 221 is the determined cDNA sequence for g-f12fwd  
 20 SEQ ID NO: 222 is the determined cDNA sequence for g-f3rev  
 SEQ ID NO: 223 is the determined cDNA sequence for P509S  
 SEQ ID NO: 224 is the determined cDNA sequence for P510S  
 SEQ ID NO: 225 is the determined cDNA sequence for P703DE5  
 SEQ ID NO: 226 is the determined cDNA sequence for 9-A11  
 25 SEQ ID NO: 227 is the determined cDNA sequence for 8-C6  
 SEQ ID NO: 228 is the determined cDNA sequence for 8-H7  
 SEQ ID NO: 229 is the determined cDNA sequence for JPTPN13  
 SEQ ID NO: 230 is the determined cDNA sequence for JPTPN14  
 SEQ ID NO: 231 is the determined cDNA sequence for JPTPN23



SEQ ID NO: 232 is the determined cDNA sequence for JPTPN24  
 SEQ ID NO: 233 is the determined cDNA sequence for JPTPN25  
 SEQ ID NO: 234 is the determined cDNA sequence for JPTPN30  
 SEQ ID NO: 235 is the determined cDNA sequence for JPTPN34  
 5 SEQ ID NO: 236 is the determined cDNA sequence for PTPN35  
 SEQ ID NO: 237 is the determined cDNA sequence for JPTPN36  
 SEQ ID NO: 238 is the determined cDNA sequence for JPTPN38  
 SEQ ID NO: 239 is the determined cDNA sequence for JPTPN39  
 SEQ ID NO: 240 is the determined cDNA sequence for JPTPN40  
 10 SEQ ID NO: 241 is the determined cDNA sequence for JPTPN41  
 SEQ ID NO: 242 is the determined cDNA sequence for JPTPN42  
 SEQ ID NO: 243 is the determined cDNA sequence for JPTPN45  
 SEQ ID NO: 244 is the determined cDNA sequence for JPTPN46  
 SEQ ID NO: 245 is the determined cDNA sequence for JPTPN51  
 15 SEQ ID NO: 246 is the determined cDNA sequence for JPTPN56  
 SEQ ID NO: 247 is the determined cDNA sequence for PTPN64  
 SEQ ID NO: 248 is the determined cDNA sequence for JPTPN65  
 SEQ ID NO: 249 is the determined cDNA sequence for JPTPN67  
 SEQ ID NO: 250 is the determined cDNA sequence for JPTPN76  
 20 SEQ ID NO: 251 is the determined cDNA sequence for JPTPN84  
 SEQ ID NO: 252 is the determined cDNA sequence for JPTPN85  
 SEQ ID NO: 253 is the determined cDNA sequence for JPTPN86  
 SEQ ID NO: 254 is the determined cDNA sequence for JPTPN87  
 SEQ ID NO: 255 is the determined cDNA sequence for JPTPN88  
 25 SEQ ID NO: 256 is the determined cDNA sequence for JP1F1  
 SEQ ID NO: 257 is the determined cDNA sequence for JP1F2  
 SEQ ID NO: 258 is the determined cDNA sequence for JP1C2  
 SEQ ID NO: 259 is the determined cDNA sequence for JP1B1  
 SEQ ID NO: 260 is the determined cDNA sequence for JP1B2



SEQ ID NO: 261 is the determined cDNA sequence for JP1D3  
 SEQ ID NO: 262 is the determined cDNA sequence for JP1A4  
 SEQ ID NO: 263 is the determined cDNA sequence for JP1F5  
 SEQ ID NO: 264 is the determined cDNA sequence for JP1E6  
 5 SEQ ID NO: 265 is the determined cDNA sequence for JP1D6  
 SEQ ID NO: 266 is the determined cDNA sequence for JP1B5  
 SEQ ID NO: 267 is the determined cDNA sequence for JP1A6  
 SEQ ID NO: 268 is the determined cDNA sequence for JP1E8  
 SEQ ID NO: 269 is the determined cDNA sequence for JP1D7  
 10 SEQ ID NO: 270 is the determined cDNA sequence for JP1D9  
 SEQ ID NO: 271 is the determined cDNA sequence for JP1C10  
 SEQ ID NO: 272 is the determined cDNA sequence for JP1A9  
 SEQ ID NO: 273 is the determined cDNA sequence for JP1F12  
 SEQ ID NO: 274 is the determined cDNA sequence for JP1E12  
 15 SEQ ID NO: 275 is the determined cDNA sequence for JP1D11  
 SEQ ID NO: 276 is the determined cDNA sequence for JP1C11  
 SEQ ID NO: 277 is the determined cDNA sequence for JP1C12  
 SEQ ID NO: 278 is the determined cDNA sequence for JP1B12  
 SEQ ID NO: 279 is the determined cDNA sequence for JP1A12  
 20 SEQ ID NO: 280 is the determined cDNA sequence for JP8G2  
 SEQ ID NO: 281 is the determined cDNA sequence for JP8H1  
 SEQ ID NO: 282 is the determined cDNA sequence for JP8H2  
 SEQ ID NO: 283 is the determined cDNA sequence for JP8A3  
 SEQ ID NO: 284 is the determined cDNA sequence for JP8A4  
 25 SEQ ID NO: 285 is the determined cDNA sequence for JP8C3  
 SEQ ID NO: 286 is the determined cDNA sequence for JP8G4  
 SEQ ID NO: 287 is the determined cDNA sequence for JP8B6  
 SEQ ID NO: 288 is the determined cDNA sequence for JP8D6  
 SEQ ID NO: 289 is the determined cDNA sequence for JP8F5



SEQ ID NO: 290 is the determined cDNA sequence for JP8A8  
 SEQ ID NO: 291 is the determined cDNA sequence for JP8C7  
 SEQ ID NO: 292 is the determined cDNA sequence for JP8D7  
 SEQ ID NO: 293 is the determined cDNA sequence for P8D8  
 5 SEQ ID NO: 294 is the determined cDNA sequence for JP8E7  
 SEQ ID NO: 295 is the determined cDNA sequence for JP8F8  
 SEQ ID NO: 296 is the determined cDNA sequence for JP8G8  
 SEQ ID NO: 297 is the determined cDNA sequence for JP8B10  
 SEQ ID NO: 298 is the determined cDNA sequence for JP8C10  
 10 SEQ ID NO: 299 is the determined cDNA sequence for JP8E9  
 SEQ ID NO: 300 is the determined cDNA sequence for JP8E10  
 SEQ ID NO: 301 is the determined cDNA sequence for JP8F9  
 SEQ ID NO: 302 is the determined cDNA sequence for JP8H9  
 SEQ ID NO: 303 is the determined cDNA sequence for JP8C12  
 15 SEQ ID NO: 304 is the determined cDNA sequence for JP8E11  
 SEQ ID NO: 305 is the determined cDNA sequence for JP8E12  
 SEQ ID NO: 306 is the amino acid sequence for the peptide PS2#12  
 SEQ ID NO: 307 is the determined cDNA sequence for P711P  
 SEQ ID NO: 308 is the determined cDNA sequence for P712P  
 20 SEQ ID NO: 309 is the determined cDNA sequence for CLONE23  
 SEQ ID NO: 310 is the determined cDNA sequence for P774P  
 SEQ ID NO: 311 is the determined cDNA sequence for P775P  
 SEQ ID NO: 312 is the determined cDNA sequence for P715P  
 SEQ ID NO: 313 is the determined cDNA sequence for P710P  
 25 SEQ ID NO: 314 is the determined cDNA sequence for P767P  
 SEQ ID NO: 315 is the determined cDNA sequence for P768P  
 SEQ ID NO: 316-325 are the determined cDNA sequences of previously  
 isolated genes  
 SEQ ID NO: 326 is the determined cDNA sequence for P703PDE5



- SEQ ID NO: 327 is the predicted amino acid sequence for P703PDE5  
 SEQ ID NO: 328 is the determined cDNA sequence for P703P6.26  
 SEQ ID NO: 329 is the predicted amino acid sequence for P703P6.26  
 SEQ ID NO: 330 is the determined cDNA sequence for P703PX-23  
 5 SEQ ID NO: 331 is the predicted amino acid sequence for P703PX-23  
 SEQ ID NO: 332 is the determined full length cDNA sequence for P509S  
 SEQ ID NO: 333 is the determined extended cDNA sequence for P707P  
 (also referred to as 11-C9)  
 SEQ ID NO: 334 is the determined cDNA sequence for P714P  
 10 SEQ ID NO: 335 is the determined cDNA sequence for P705P (also  
 referred to as 9-F3)  
 SEQ ID NO: 336 is the predicted amino acid sequence for P705P  
 SEQ ID NO: 337 is the amino acid sequence of the peptide P1S#10  
 SEQ ID NO: 338 is the amino acid sequence of the peptide p5  
 15 SEQ ID NO: 339 is the predicted amino acid sequence of P509S  
 SEQ ID NO: 340 is the determined cDNA sequence for P778P  
 SEQ ID NO: 341 is the determined cDNA sequence for P786P  
 SEQ ID NO: 342 is the determined cDNA sequence for P789P  
 SEQ ID NO: 343 is the determined cDNA sequence for a clone showing  
 20 homology to Homo sapiens MM46 mRNA  
 SEQ ID NO: 344 is the determined cDNA sequence for a clone showing  
 homology to Homo sapiens TNF-alpha stimulated ABC protein (ABC50) mRNA  
 SEQ ID NO: 345 is the determined cDNA sequence for a clone showing  
 homology to Homo sapiens mRNA for E-cadherin  
 25 SEQ ID NO: 346 is the determined cDNA sequence for a clone showing  
 homology to Human nuclear-encoded mitochondrial serine hydroxymethyltransferase  
 (SHMT)  
 SEQ ID NO: 347 is the determined cDNA sequence for a clone showing  
 homology to Homo sapiens natural resistance-associated macrophage protein2 (NRAMP2)



SEQ ID NO: 348 is the determined cDNA sequence for a clone showing homology to Homo sapiens phosphoglucomutase-related protein (PGMRP)

SEQ ID NO: 349 is the determined cDNA sequence for a clone showing homology to Human mRNA for proteosome subunit p40

5 SEQ ID NO: 350 is the determined cDNA sequence for P777P

SEQ ID NO: 351 is the determined cDNA sequence for P779P

SEQ ID NO: 352 is the determined cDNA sequence for P790P

SEQ ID NO: 353 is the determined cDNA sequence for P784P

SEQ ID NO: 354 is the determined cDNA sequence for P776P

10 SEQ ID NO: 355 is the determined cDNA sequence for P780P

SEQ ID NO: 356 is the determined cDNA sequence for P544S

SEQ ID NO: 357 is the determined cDNA sequence for P745S

SEQ ID NO: 358 is the determined cDNA sequence for P782P

SEQ ID NO: 359 is the determined cDNA sequence for P783P

15 SEQ ID NO: 360 is the determined cDNA sequence for unknown 17984

SEQ ID NO: 361 is the determined cDNA sequence for P787P

SEQ ID NO: 362 is the determined cDNA sequence for P788P

SEQ ID NO: 363 is the determined cDNA sequence for unknown 17994

SEQ ID NO: 364 is the determined cDNA sequence for P781P

20 SEQ ID NO: 365 is the determined cDNA sequence for P785P

SEQ ID NO: 366-375 are the determined cDNA sequences for splice variants of B305D.

SEQ ID NO: 376 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 366.

25 SEQ ID NO: 377 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 372.

SEQ ID NO: 378 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 373.



SEQ ID NO: 379 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 374.

SEQ ID NO: 380 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 375.

- 5           SEQ ID NO: 381 is the determined cDNA sequence for B716P.  
          SEQ ID NO: 382 is the determined full-length cDNA sequence for P711P.  
          SEQ ID NO: 383 is the predicted amino acid sequence for P711P.  
          SEQ ID NO: 384 is the cDNA sequence for P1000C.  
          SEQ ID NO: 385 is the cDNA sequence for CGI-82.
- 10          SEQ ID NO:386 is the cDNA sequence for 23320.  
          SEQ ID NO:387 is the cDNA sequence for CGI-69.  
          SEQ ID NO:388 is the cDNA sequence for L-iditol-2-dehydrogenase.  
          SEQ ID NO:389 is the cDNA sequence for 23379.  
          SEQ ID NO:390 is the cDNA sequence for 23381.
- 15          SEQ ID NO:391 is the cDNA sequence for KIAA0122.  
          SEQ ID NO:392 is the cDNA sequence for 23399.  
          SEQ ID NO:393 is the cDNA sequence for a previously identified gene.  
          SEQ ID NO:394 is the cDNA sequence for HCLBP.  
          SEQ ID NO:395 is the cDNA sequence for transglutaminase.
- 20          SEQ ID NO:396 is the cDNA sequence for a previously identified gene.  
          SEQ ID NO:397 is the cDNA sequence for PAP.  
          SEQ ID NO:398 is the cDNA sequence for Ets transcription factor PDEF.  
          SEQ ID NO:399 is the cDNA sequence for hTGR.  
          SEQ ID NO:400 is the cDNA sequence for KIAA0295.
- 25          SEQ ID NO:401 is the cDNA sequence for 22545.  
          SEQ ID NO:402 is the cDNA sequence for 22547.  
          SEQ ID NO:403 is the cDNA sequence for 22548.  
          SEQ ID NO:404 is the cDNA sequence for 22550.  
          SEQ ID NO:405 is the cDNA sequence for 22551.



SEQ ID NO:406 is the cDNA sequence for 22552.

SEQ ID NO:407 is the cDNA sequence for 22553 (also known as P1020C).

SEQ ID NO:408 is the cDNA sequence for 22558.

SEQ ID NO:409 is the cDNA sequence for 22562.

5 SEQ ID NO:410 is the cDNA sequence for 22565.

SEQ ID NO:411 is the cDNA sequence for 22567.

SEQ ID NO:412 is the cDNA sequence for 22568.

SEQ ID NO:413 is the cDNA sequence for 22570.

SEQ ID NO:414 is the cDNA sequence for 22571.

10 SEQ ID NO:415 is the cDNA sequence for 22572.

SEQ ID NO:416 is the cDNA sequence for 22573.

SEQ ID NO:417 is the cDNA sequence for 22573.

SEQ ID NO:418 is the cDNA sequence for 22575.

SEQ ID NO:419 is the cDNA sequence for 22580.

15 SEQ ID NO:420 is the cDNA sequence for 22581.

SEQ ID NO:421 is the cDNA sequence for 22582.

SEQ ID NO:422 is the cDNA sequence for 22583.

SEQ ID NO:423 is the cDNA sequence for 22584.

SEQ ID NO:424 is the cDNA sequence for 22585.

20 SEQ ID NO:425 is the cDNA sequence for 22586.

SEQ ID NO:426 is the cDNA sequence for 22587.

SEQ ID NO:427 is the cDNA sequence for 22588.

SEQ ID NO:428 is the cDNA sequence for 22589.

SEQ ID NO:429 is the cDNA sequence for 22590.

25 SEQ ID NO:430 is the cDNA sequence for 22591.

SEQ ID NO:431 is the cDNA sequence for 22592.

SEQ ID NO:432 is the cDNA sequence for 22593.

SEQ ID NO:433 is the cDNA sequence for 22594.

SEQ ID NO:434 is the cDNA sequence for 22595.



SEQ ID NO:435 is the cDNA sequence for 22596.  
 SEQ ID NO:436 is the cDNA sequence for 22847.  
 SEQ ID NO:437 is the cDNA sequence for 22848.  
 SEQ ID NO:438 is the cDNA sequence for 22849.  
 5 SEQ ID NO:439 is the cDNA sequence for 22851.  
 SEQ ID NO:440 is the cDNA sequence for 22852.  
 SEQ ID NO:441 is the cDNA sequence for 22853.  
 SEQ ID NO:442 is the cDNA sequence for 22854.  
 SEQ ID NO:443 is the cDNA sequence for 22855.  
 10 SEQ ID NO:444 is the cDNA sequence for 22856.  
 SEQ ID NO:445 is the cDNA sequence for 22857.  
 SEQ ID NO:446 is the cDNA sequence for 23601.  
 SEQ ID NO:447 is the cDNA sequence for 23602.  
 SEQ ID NO:448 is the cDNA sequence for 23605.  
 15 SEQ ID NO:449 is the cDNA sequence for 23606.  
 SEQ ID NO:450 is the cDNA sequence for 23612.  
 SEQ ID NO:451 is the cDNA sequence for 23614.  
 SEQ ID NO:452 is the cDNA sequence for 23618.  
 SEQ ID NO:453 is the cDNA sequence for 23622.  
 20 SEQ ID NO:454 is the cDNA sequence for folate hydrolase.  
 SEQ ID NO:455 is the cDNA sequence for LIM protein.  
 SEQ ID NO:456 is the cDNA sequence for a known gene.  
 SEQ ID NO:457 is the cDNA sequence for a known gene.  
 SEQ ID NO:458 is the cDNA sequence for a previously identified gene.  
 25 SEQ ID NO:459 is the cDNA sequence for 23045.  
 SEQ ID NO:460 is the cDNA sequence for 23032.  
 SEQ ID NO:461 is the cDNA sequence for 23054.  
 SEQ ID NO:462-467 are cDNA sequences for known genes.  
 SEQ ID NO:468-471 are cDNA sequences for P710P.



SEQ ID NO:472 is a cDNA sequence for P1001C.

SEQ ID NO: 473 is the determined cDNA sequence for a first splice variant of P775P (referred to as 27505).

5 SEQ ID NO: 474 is the determined cDNA sequence for a second splice variant of P775P (referred to as 19947).

SEQ ID NO: 475 is the determined cDNA sequence for a third splice variant of P775P (referred to as 19941).

SEQ ID NO: 476 is the determined cDNA sequence for a fourth splice variant of P775P (referred to as 19937).

10 SEQ ID NO: 477 is a first predicted amino acid sequence encoded by the sequence of SEQ ID NO: 474.

SEQ ID NO: 478 is a second predicted amino acid sequence encoded by the sequence of SEQ ID NO: 474.

15 SEQ ID NO: 479 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 475.

SEQ ID NO: 480 is a first predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 481 is a second predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

20 SEQ ID NO: 482 is a third predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 483 is a fourth predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

25 SEQ ID NO: 484 is the first 30 amino acids of the *M. tuberculosis* antigen Ra12.

SEQ ID NO: 485 is the PCR primer AW025.

SEQ ID NO: 486 is the PCR primer AW003.

SEQ ID NO: 487 is the PCR primer AW027.

SEQ ID NO: 488 is the PCR primer AW026.



SEQ ID NO: 502 is the determined cDNA sequence of the complementarity determining region for the anti-P503S monoclonal antibody 20D4.

SEQ ID NO: 503 is the determined cDNA sequence of the complementarity determining region for the anti-P503S monoclonal antibody JA1.

SEQ ID NO: 504 & 505 are peptides employed in epitope mapping studies.

SEQ ID NO: 506 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 8H2.

SEQ ID NO: 507 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 7H8.

SEQ ID NO: 508 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 2D4.

SEQ ID NO: 509-522 are peptides employed in epitope mapping studies.

SEQ ID NO: 523 is a mature form of P703P used to raise antibodies against P703P.

SEQ ID NO: 524 is the putative full-length cDNA sequence of P703P.

SEQ ID NO: 525 is the predicted amino acid sequence encoded by SEQ ID NO: 524.

SEQ ID NO: 526 is the full-length cDNA sequence for P790P.

SEQ ID NO: 527 is the predicted amino acid sequence for P790P.

SEQ ID NO: 528 & 529 are PCR primers.

SEQ ID NO: 530 is the cDNA sequence of a splice variant of SEQ ID NO: 366.

SEQ ID NO: 531 is the cDNA sequence of the open reading frame of SEQ ID NO: 530.

SEQ ID NO: 532 is the predicted amino acid encoded by the sequence of SEQ ID NO: 531.

SEQ ID NO: 533 is the DNA sequence of a putative ORF of P775P.



SEQ ID NO: 534 is the predicted amino acid sequence encoded by SEQ ID NO: 533.

SEQ ID NO: 535 is a first full-length cDNA sequence for P510S.

SEQ ID NO: 536 is a second full-length cDNA sequence for P510S.

5 SEQ ID NO: 537 is the predicted amino acid sequence encoded by SEQ ID NO: 535.

SEQ ID NO: 538 is the predicted amino acid sequence encoded by SEQ ID NO: 536.

SEQ ID NO: 539 is the peptide P501S-370.

10 SEQ ID NO: 540 is the peptide P501S-376.

SEQ ID NO: 541-551 are epitopes of P501S.

SEQ ID NO: 552 is an extended cDNA sequence for P712P.

SEQ ID NO: 553-568 are the amino acid sequences encoded by predicted open reading frames within SEQ ID NO: 552.

15 SEQ ID NO: 569 is an extended cDNA sequence for P776P.

SEQ ID NO: 570 is the determined cDNA sequence for a splice variant of P776P referred to as contig 6.

SEQ ID NO: 571 is the determined cDNA sequence for a splice variant of P776P referred to as contig 7.

20 SEQ ID NO: 572 is the determined cDNA sequence for a splice variant of P776P referred to as contig 14.

SEQ ID NO: 573 is the amino acid sequence encoded by a first predicted ORF of SEQ ID NO: 570.

25 SEQ ID NO: 574 is the amino acid sequence encoded by a second predicted ORF of SEQ ID NO: 570.

SEQ ID NO: 575 is the amino acid sequence encoded by a predicted ORF of SEQ ID NO: 571.

SEQ ID NO: 576-586 are amino acid sequences encoded by predicted ORFs of SEQ ID NO: 569.



SEQ ID NO: 587 is a DNA consensus sequence of the sequences of P767P and P777P.

SEQ ID NO: 588-590 are amino acid sequences encoded by predicted ORFs of SEQ ID NO: 587.

5 SEQ ID NO: 591 is an extended cDNA sequence for P1020C.

SEQ ID NO: 592 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: P1020C.

SEQ ID NO: 593 is a splice variant of P775P referred to as 50748.

10 SEQ ID NO: 594 is a splice variant of P775P referred to as 50717. SEQ ID NO: 595 is a splice variant of P775P referred to as 45985.

SEQ ID NO: 596 is a splice variant of P775P referred to as 38769.

SEQ ID NO: 597 is a splice variant of P775P referred to as 37922.

SEQ ID NO: 598 is a splice variant of P510S referred to as 49274.

SEQ ID NO: 599 is a splice variant of P510S referred to as 39487.

15 SEQ ID NO: 600 is a splice variant of P504S referred to as 5167.16.

SEQ ID NO: 601 is a splice variant of P504S referred to as 5167.1.

SEQ ID NO: 602 is a splice variant of P504S referred to as 5163.46.

SEQ ID NO: 603 is a splice variant of P504S referred to as 5163.42.

SEQ ID NO: 604 is a splice variant of P504S referred to as 5163.34.

20 SEQ ID NO: 605 is a splice variant of P504S referred to as 5163.17.

SEQ ID NO: 606 is a splice variant of P501S referred to as 10640.

SEQ ID NO: 607-615 are the sequences of PCR primers.

SEQ ID NO: 616 is the determined cDNA sequence of a fusion of P703P and PSA.

25 SEQ ID NO: 617 is the amino acid sequence of the fusion of P703P and PSA.

SEQ ID NO: 618-689 are determined cDNA sequences of prostate-specific clones.

SEQ ID NO: 690 is the cDNA sequence of the gene DD3.



SEQ ID NO: 691-697 are determined cDNA sequences of prostate-specific clones.

SEQ ID NO: 698 is an extended cDNA sequence for P714P.

SEQ ID NO: 699-701 are the cDNA sequences for splice variants of P704P.

5 SEQ ID NO: 702 is the cDNA sequence of a spliced variant of P553S referred to as P553S-14.

SEQ ID NO: 703 is the cDNA sequence of a spliced variant of P553S referred to as P553S-12.

10 SEQ ID NO: 704 is the cDNA sequence of a spliced variant of P553S referred to as P553S-10.

SEQ ID NO: 705 is the cDNA sequence of a spliced variant of P553S referred to as P553S-6.

SEQ ID NO: 706 is the amino acid sequence encoded by SEQ ID NO: 705.

SEQ ID NO: 707 is the amino acid sequence encoded by SEQ ID NO: 702

15 SEQ ID NO: 708 is a second amino acid sequence encoded by SEQ ID NO: 702.

SEQ ID NO: 709-772 are determined cDNA sequences of prostate-specific clones.

SEQ ID NO: 773 is a first full-length cDNA sequence for prostate-specific transglutaminase gene (also referred to herein as P558S).

20 SEQ ID NO: 774 is a second full-length cDNA sequence for prostate-specific transglutaminase gene.

SEQ ID NO: 775 is the amino acid sequence encoded by the sequence of SEQ ID NO: 773.

25 SEQ ID NO: 776 is the amino acid sequence encoded by the sequence of SEQ ID NO: 774.

SEQ ID NO: 777 is the full-length cDNA sequence for P788P.

SEQ ID NO: 778 is the amino acid sequence encoded by SEQ ID NO: 777.

SEQ ID NO: 779 is the determined cDNA sequence for a polymorphic variant of P788P.



SEQ ID NO: 780 is the amino acid sequence encoded by SEQ ID NO: 779.

SEQ ID NO: 781 is the amino acid sequence of peptide 4 from P703P.

SEQ ID NO: 782 is the cDNA sequence that encodes peptide 4 from P703P.

SEQ ID NO: 783-798 are the cDNA sequence encoding epitopes of P703P.

5 SEQ ID NO: 799-814 are the amino acid sequences of epitopes of P703P.

## DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for using the compositions, for example in the therapy and diagnosis of cancer, such as prostate cancer. Certain illustrative compositions described herein include  
 10 prostate-specific polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (*e.g.*, T cells). A "prostate-specific protein," as the term is used herein, refers generally to a protein that is expressed in prostate cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in other normal tissues, as determined using a  
 15 representative assay provided herein. Certain prostate-specific proteins are tumor proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with prostate cancer.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set  
 20 forth in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789, illustrative polypeptide compositions having amino acid sequences set forth in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534,  
 25 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778 and 780, antibody compositions capable of binding such polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human prostate cancer.



## POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains  
 5 one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

10 As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

15 "Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the  
 20 hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules,  
 25 which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a prostate-specific protein or a portion thereof) or may comprise a



variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the  
 5 immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term “variants” also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be “identical” if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons  
 10 between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A “comparison window” as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences  
 15 are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of  
 20 evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenesis pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS*  
 25 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.



Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the "percentage of sequence identity" is determined by comparing two optimally aligned sequences over a window of comparison of at least 20



positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.



The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length  
 5 may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like,  
 10 (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof.  
 15 Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X  
 20 and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that  
 25 vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an



altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

### PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences  
 5 provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary  
 10 sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of  
 15 complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so  
 20 (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately  
 25 depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100



nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having  
 5 contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

10 Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789, or to any continuous portion of the sequence,  
 15 from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for  
 20 example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR<sup>TM</sup> technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA  
 25 techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of



probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about  
 5 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less  
 10 stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any  
 15 case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

## 20 POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using any of a variety of well established techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than  
 25 in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller *et al.*, *Proc. Natl. Acad. Sci. USA*



94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as prostate-specific cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided  
 5 herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.*, a prostate tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for  
 10 amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (*e.g.*, by nick-translation or end-labeling with  $^{32}\text{P}$ ) using well known techniques. A bacterial or  
 15 bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be  
 20 analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then  
 25 assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available



kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and  
 5 overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia *et al.*, *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within  
 10 an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from  
 15 the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom *et al.*, *PCR Methods Applic.* 1:111-19, 1991)  
 20 and walking PCR (Parker *et al.*, *Nucl. Acids. Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed  
 25 using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.



## POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct  
5 expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in  
10 some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be  
15 engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic  
20 oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant  
25 nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site



located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. *et al.* (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.



A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems  
 5 infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression  
 10 vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when  
 15 cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide,  
 20 vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct  
 25 high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced;



pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel *et al.* (supra) and Grant *et al.* (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. *et al.* (1984) *EMBO J.* 3:1671-1680; Broglie, R. *et al.* (1984) *Science* 224:838-843; and Winter, J. *et al.* (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the



polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired



fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 5 HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a 10 polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, 15 and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase 20 (Wigler, M. *et al.* (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the 25 aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C.



Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein  
 5 expression attributable to a specific vector system (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence,  
 10 recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

15 Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

20 A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal  
 25 antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. *et al.* (1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) and Maddox, D. E. *et al.* (1983; *J. Exp. Med.* 158:1211-1216).



A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a



nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. *et al.* (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

#### SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.



In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to  
 5 create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

10 As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also  
 15 routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the  
 20 desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and  
 25 the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful



species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term “oligonucleotide directed mutagenesis procedure” refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term “oligonucleotide directed mutagenesis procedure” is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

#### **POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES**

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR™) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCR™, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An



excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR™, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α-thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated



herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.



Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to



make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

5 PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other  
10 amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in  
15 its entirety), may also be used in the amplification of DNA sequences of the present invention.

#### **BIOLOGICAL FUNCTIONAL EQUIVALENTS**

Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional  
20 molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

25 When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.



For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.



**TABLE 1**

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, 5 incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been



assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within  $\pm 2$  is preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0  $\pm$  1); glutamate (+3.0  $\pm$  1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (−0.4); proline (−0.5  $\pm$  1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within  $\pm 2$  is preferred, those within  $\pm 1$  are particularly preferred, and those within  $\pm 0.5$  are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their



hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

5           In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl- methyl-, thio- and other  
10   modified forms of adenine, cytidine, guanine, thymine and uridine.

#### **IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES**

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety of well known approaches, several of which are outlined below for  
15   the purpose of illustration.

##### **1.     ADENOVIRUS**

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences  
20   sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-  
25   stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because



adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease  
 5 such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The  
 10 early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication,  
 15 late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-tripartite leader (TPL) sequence  
 20 which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual  
 25 plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is



dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.



Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup  
 5 C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is  
 10 replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted  
 15 E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*,  $10^9$ - $10^{11}$  plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not  
 20 require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

25 Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant



adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

## 5 2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results  
10 in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome.  
15 These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order  
20 to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be  
25 packaged into viral particles, which are then secreted into the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene



transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical  
5 addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using  
10 streptavidin (Roux *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

### 3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzyczka, 1984) is a parovirus,  
15 discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid  
20 proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs. There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications,  
25 whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral



promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

#### 4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive



properties for liver-directed gene transfer. Chang *et al.* (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

## 5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell



membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e. ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present invention.

#### ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for



polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism  
 5 to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense  
 10 inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA<sub>A</sub> receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense  
 15 constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides  
 20 oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a  
 25 phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.



Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the rat and human sequences) and determination of secondary structure,  $T_m$ , binding energy, relative stability, and antisense compositions were selected based upon their relative  
 5 inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target  
 10 site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic  
 15 domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma  
 20 membrane (Morris *et al.*, 1997).

## RIBOZYMES

Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes  
 25 have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et*



*al.*, 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

5               Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, 10               sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes *H-ras*, *c-fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon 15               that is cleaved by a specific ribozyme.

              Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through 20               the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an 25               encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

              The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to



a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the  
 5 ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, 1992). Thus, the specificity of action  
 10 of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis  $\delta$  virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are  
 15 described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis  $\delta$  virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA  
 20 ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S. Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it  
 25 have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as



one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific  
 5 cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells  
 10 from eukaryotic promoters (*e.g.*, Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No.  
 15 WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo*  
 20 through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such  
 25 ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into



the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target  
5 RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid  
10 protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-  
15 C-allyl, 2'-fluoro, 2'-o-methyl, 2'-H (for a review see *e.g.*, Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their  
20 degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al.*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of  
25 enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be



administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisiewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).



Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

#### PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter,



decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and  
 5 Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a  
 10 modified Merrifield method, have been used (Christensen *et al.*, 1995).

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, 1995). The manual protocol lends itself to the production of chemically modified PNAs or the  
 15 simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is  
 20 suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

25 Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific



functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaime *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ulmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremsky *et al.*, 1996; Pardridge *et al.*, 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

10 In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3' to 5' polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

15 Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature ( $T_m$ ) and reduces the dependence of  $T_m$  on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The enhanced rate and affinity of hybridization are significant because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang *et al.*, 1996).

25 One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence



specificity of PNA recognition. As with DNA hybridization, however, selective recognition can be achieved by balancing oligomer length and incubation temperature. Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a  
 5 single mismatch within a 16 bp PNA-DNA duplex can reduce the  $T_m$  by up to 15°C (Egholm *et al.*, 1993). This high level of discrimination has allowed the development of several PNA-based strategies for the analysis of point mutations (Wang *et al.*, 1996; Carlsson *et al.*, 1996; Thiede *et al.*, 1996; Webb and Hurskainen, 1996; Perry-O'Keefe *et al.*, 1996).

10 High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an essential RNA template, while the analogous DNA oligomers do not (Norton *et al.*, 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and  
 15 this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini (Nielsen *et al.*, 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur  
 20 spontaneously at sequences within chromosomal DNA (Boffa *et al.*, 1995; Boffa *et al.*, 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa *et al.*, 1995) and to inhibit transcription (Boffa *et al.*, 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene  
 25 expression. Studies and reviews concerning the use of PNAs as antisense and anti-gene agents include Nielsen *et al.* (1993b), Hanvey *et al.* (1992), and Good and Nielsen (1997). Koppelhus *et al.* (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.



Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by  
 5 Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995), blocking of  
 10 transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

#### **POLYPEPTIDE COMPOSITIONS**

The present invention, in other aspects, provides polypeptide compositions.  
 15 Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a  
 20 contiguous amino acid sequence from an amino acid sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, particularly a polypeptide having the  
 25 amino acid sequence disclosed in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778 and 780, or active fragments, variants or biological functional equivalents thereof.



Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789, or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency. Particularly illustrative polypeptides include the amino acid sequence disclosed in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778 and 780.

As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of a prostate-specific protein or a variant thereof, as described herein. As noted above, a "prostate-specific protein" is a protein that is expressed by prostate cells. Proteins that are prostate-specific proteins also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with prostate cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a prostate-specific protein or a variant thereof. Certain preferred immunogenic portions



include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

5                   Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they  
10 specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native prostate-specific protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full  
15 length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For  
20 example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, <sup>125</sup>I-labeled Protein A.

As noted above, a composition may comprise a variant of a native prostate-specific protein. A polypeptide "variant," as used herein, is a polypeptide that differs from  
25 a native prostate-specific protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than



50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more  
 5 portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those  
 10 exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has  
 15 similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and  
 20 glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val,  
 25 ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the



deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydrophobic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from



suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of



hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*, *Gene* 40:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

10           The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

15           Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute *et al.* *New Engl. J. Med.*, 336:86-91, 1997).

20           Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a  
25 Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein



from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

## **BINDING AGENTS**

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a prostate-specific protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a prostate-specific protein if it reacts at a detectable level (within, for example, an



ELISA) with a prostate-specific protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about  $10^3$  L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as prostate cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to a prostate-specific protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of



monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the



yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and  
 5 extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit  
 10 serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or  
 15 more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include  $^{90}\text{Y}$ ,  $^{123}\text{I}$ ,  $^{125}\text{I}$ ,  $^{131}\text{I}$ ,  $^{186}\text{Re}$ ,  $^{188}\text{Re}$ ,  $^{211}\text{At}$ , and  $^{212}\text{Bi}$ . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria  
 20 toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent  
 25 capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.



Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter *et al.*), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn *et al.*), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler *et al.*).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled



directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato *et al.*), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih *et al.*). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison *et al.* discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

## 20 T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a prostate-specific protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243).



Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a prostate-specific polypeptide, polynucleotide encoding a prostate-specific polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a prostate-specific polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a prostate-specific polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen *et al.*, *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a prostate-specific polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan *et al.*, *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a prostate-specific polypeptide, polynucleotide or polypeptide-expressing APC may be CD4<sup>+</sup> and/or CD8<sup>+</sup>. prostate-specific protein-specific T cells may be expanded using standard techniques.



Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4<sup>+</sup> or CD8<sup>+</sup> T cells that proliferate in response to a prostate-specific polypeptide, polynucleotide or APC can be expanded in number  
 5 either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a prostate-specific polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a prostate-specific polypeptide. Alternatively, one or  
 10 more T cells that proliferate in the presence of a prostate-specific protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

#### PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of  
 15 one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be  
 20 administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the  
 25 particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA compositions.



Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

## 1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup or elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition,



the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

## 2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety).



Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations  
 5 contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the  
 10 extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper  
 15 fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic  
 20 agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered  
 25 isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or



injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption



delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active  
 5 ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions  
 10 are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

### 3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered  
 15 by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-  
 20 glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

### 4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

25 In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In



particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of  
 5 pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved  
 10 serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent  
 15 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In  
 20 addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric  
 25 effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of



liposomes is not associated with autoimmune responses, toxicity or gonadal localization after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)). MLVs generally have diameters of from 25 nm to 4  $\mu$ m. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur *et al.* (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the



most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains



the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In  
5 general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be  
10 accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell  
15 types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*,  
20 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1  $\mu\text{m}$ ) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be are easily made, as described (Couvreur *et al.*,  
25 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by reference in its entirety).



## IMMUNOGENIC COMPOSITIONS

In certain preferred embodiments of the present invention, immunogenic compositions, or vaccines, are provided. The immunogenic compositions will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and immunogenic compositions within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition.

Illustrative immunogenic compositions may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for



example, in Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner *et al.*, *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner *et al.*, *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 5 6:616-627, 1988; Rosenfeld *et al.*, *Science* 252:431-434, 1991; Kolls *et al.*, *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman *et al.*, *Circulation* 88:2838-2848, 1993; and Guzman *et al.*, *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be 10 "naked," as described, for example, in Ulmer *et al.*, *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that an immunogenic composition may comprise both a polynucleotide and a polypeptide component. Such immunogenic compositions may 15 provide for an enhanced immune response.

It will be apparent that an immunogenic composition may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) 20 and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for 25 any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate,



sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the immunogenic compositions of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.



Within the immunogenic compositions provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN- $\gamma$ , TNF $\alpha$ , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of an immunogenic composition as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science* 273:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.



Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and  
5 other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

Any immunogenic composition provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a  
10 suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine* 14:1429-1438, 1996) and administered by,  
15 for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also  
20 be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer  
25 comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.



Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and immunogenic compositions to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within an immunogenic composition (*see* Zitvogel *et al.*, *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4,



IL-13 and/or TNF $\alpha$  to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF $\alpha$ , CD40 ligand, LPS, flt3 ligand and/or other  
 5 compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC  
 10 with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc $\gamma$  receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86  
 15 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a prostate-specific protein (or portion or other variant thereof) such that the prostate-specific polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition comprising such transfected cells  
 20 may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach  
 25 described by Mahvi *et al.*, *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the prostate-specific polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently



conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Immunogenic compositions and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a immunogenic composition or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

#### CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as prostate cancer. Within such methods, pharmaceutical compositions and immunogenic compositions are typically administered to a patient. As used herein, a “patient” refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and immunogenic compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and immunogenic compositions may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host



immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8<sup>+</sup> cytotoxic T lymphocytes and CD4<sup>+</sup> T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*.



Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see, for example, Cheever et al., Immunological Reviews 157:177, 1997*).

5                   Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

10                   Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and immunogenic compositions may be administered by injection (*e.g., intracutaneous, intramuscular, intravenous or subcutaneous*), intranasally (*e.g., by aspiration*) or orally.

15                   Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50%

20                   above the basal (*i.e., untreated*) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such immunogenic compositions should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g., more frequent remissions, complete or partial or longer disease-free*

25                   survival) in treated patients as compared to non-treated patients. In general, for pharmaceutical compositions and immunogenic compositions comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25  $\mu$ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.



In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a prostate-specific protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

## 10 **CANCER DETECTION AND DIAGNOSIS**

In general, a cancer may be detected in a patient based on the presence of one or more prostate-specific proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as prostate cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a prostate-specific sequence should be present at a level that is at least three fold higher in prostate tissue than in other normal tissues.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.



In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex.

5 Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent  
10 with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length prostate-specific proteins and portions thereof to which the binding agent binds, as described above.

15 The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic  
20 particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which  
25 may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1



hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10  $\mu$ g, and preferably about 100 ng to about 1  $\mu$ g, is sufficient to immobilize an adequate amount of binding agent.

5 Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group  
10 on the support with an amine and an active hydrogen on the binding partner (*see, e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that  
15 polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the  
20 specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20<sup>TM</sup> (Sigma Chemical Co., St. Louis, MO). The immobilized  
25 antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with prostate cancer. Preferably, the contact time is



sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as prostate cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver



Operator Curve, according to the method of Sackett *et al.*, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible

5 cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right,

10 to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the

15 immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the

20 sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result.

25 In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of



antibody immobilized on the membrane ranges from about 25 ng to about 1 $\mu$ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use prostate-specific polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such prostate-specific protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a prostate-specific protein in a biological sample. Within certain methods, a biological sample comprising CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient is incubated with a prostate-specific polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25  $\mu$ g/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of prostate-specific polypeptide to serve as a control. For CD4<sup>+</sup> T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8<sup>+</sup> T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a prostate-specific protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction



(PCR) based assay to amplify a portion of a prostate-specific cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the prostate-specific protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a prostate-specific protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a prostate-specific protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is



not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

5                   In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed  
10 as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor.  
15 One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple prostate-specific protein  
20 markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins  
25 provided herein may be combined with assays for other known tumor antigens.

## DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for



performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a prostate-specific protein. Such antibodies or fragments may be provided attached to a support material, as described  
 5 above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a  
 10 prostate-specific protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a prostate-specific protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to  
 15 facilitate the detection of a polynucleotide encoding a prostate-specific protein.

The following Examples are offered by way of illustration and not by way of limitation.



## EXAMPLE 1

## ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES

This Example describes the isolation of certain prostate-specific  
 5 polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library was constructed from prostate tumor poly A<sup>+</sup> RNA using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD 20897) following the manufacturer's protocol. Specifically, prostate tumor tissues were homogenized with  
 10 polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A<sup>+</sup> RNA was then purified using a Qiagen oligotex spin column mRNA purification kit (Qiagen, Santa Clarita, CA 91355) according to the manufacturer's protocol. First-strand cDNA was synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized,  
 15 ligated with EcoRI/BAXI adaptors (Invitrogen, San Diego, CA) and digested with NotI. Following size fractionation with Chroma Spin-1000 columns (Clontech, Palo Alto, CA), the cDNA was ligated into the EcoRI/NotI site of pCDNA3.1 (Invitrogen) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

Using the same procedure, a normal human pancreas cDNA expression  
 20 library was prepared from a pool of six tissue specimens (Clontech). The cDNA libraries were characterized by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis. The prostate tumor library contained  $1.64 \times 10^7$  independent colonies, with 70% of clones having an insert and the average insert size being 1745 base pairs. The normal pancreas cDNA  
 25 library contained  $3.3 \times 10^6$  independent colonies, with 69% of clones having inserts and the average insert size being 1120 base pairs. For both libraries, sequence analysis showed that the majority of clones had a full length cDNA sequence and were synthesized from mRNA, with minimal rRNA and mitochondrial DNA contamination.



cDNA library subtraction was performed using the above prostate tumor and normal pancreas cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a prostate tumor-specific subtracted cDNA library was generated as follows. Normal pancreas cDNA library (70 µg) was digested with  
 5 EcoRI, NotI, and SfuI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 100 µl of H<sub>2</sub>O, heat-denatured and mixed with 100 µl (100 µg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional  
 10 Photoprobe biotin (50 µl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 µl H<sub>2</sub>O to form the driver DNA.

To form the tracer DNA, 10 µg prostate tumor cDNA library was digested with BamHI and XhoI, phenol chloroform extracted and passed through Chroma spin-400  
 15 columns (Clontech). Following ethanol precipitation, the tracer DNA was dissolved in 5 µl H<sub>2</sub>O. Tracer DNA was mixed with 15 µl driver DNA and 20 µl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and incubated for 20 hours (long hybridization [LH]). The reaction  
 20 mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 µl H<sub>2</sub>O, mixed with 8 µl driver DNA and 20 µl of 2 x hybridization buffer, and subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into  
 25 BamHI/XhoI site of chloramphenicol resistant pBCSK<sup>+</sup> (Stratagene, La Jolla, CA 92037) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a prostate tumor specific subtracted cDNA library (referred to as “prostate subtraction 1”).

To analyze the subtracted cDNA library, plasmid DNA was prepared from 100 independent clones, randomly picked from the subtracted prostate tumor specific



library and grouped based on insert size. Representative cDNA clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A (Foster City, CA). Six cDNA clones, hereinafter referred to as F1-13, F1-12, F1-16, H1-1, H1-9 and H1-4, were shown to be abundant in the subtracted prostate-specific cDNA library. The determined 3' and 5' cDNA sequences for F1-12 are provided in SEQ ID NO: 2 and 3, respectively, with determined 3' cDNA sequences for F1-13, F1-16, H1-1, H1-9 and H1-4 being provided in SEQ ID NO: 1 and 4-7, respectively.

The cDNA sequences for the isolated clones were compared to known sequences in the gene bank using the EMBL and GenBank databases (release 96). Four of the prostate tumor cDNA clones, F1-13, F1-16, H1-1, and H1-4, were determined to encode the following previously identified proteins: prostate specific antigen (PSA), human glandular kallikrein, human tumor expression enhanced gene, and mitochondria cytochrome C oxidase subunit II. H1-9 was found to be identical to a previously identified human autonomously replicating sequence. No significant homologies to the cDNA sequence for F1-12 were found.

Subsequent studies led to the isolation of a full-length cDNA sequence for F1-12 (also referred to as P504S). This sequence is provided in SEQ ID NO: 107, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 108. cDNA splice variants of P504S are provided in SEQ ID NO: 600-605.

To clone less abundant prostate tumor specific genes, cDNA library subtraction was performed by subtracting the prostate tumor cDNA library described above with the normal pancreas cDNA library and with the three most abundant genes in the previously subtracted prostate tumor specific cDNA library: human glandular kallikrein, prostate specific antigen (PSA), and mitochondria cytochrome C oxidase subunit II. Specifically, 1  $\mu$ g each of human glandular kallikrein, PSA and mitochondria cytochrome C oxidase subunit II cDNAs in pCDNA3.1 were added to the driver DNA and subtraction was performed as described above to provide a second subtracted cDNA library hereinafter referred to as the "subtracted prostate tumor specific cDNA library with spike".



Twenty-two cDNA clones were isolated from the subtracted prostate tumor specific cDNA library with spike. The determined 3' and 5' cDNA sequences for the clones referred to as J1-17, L1-12, N1-1862, J1-13, J1-19, J1-25, J1-24, K1-58, K1-63, L1-4 and L1-14 are provided in SEQ ID NOS: 8-9, 10-11, 12-13, 14-15, 16-17, 18-19, 20-21, 22-23, 24-25, 26-27 and 28-29, respectively. The determined 3' cDNA sequences for the clones referred to as J1-12, J1-16, J1-21, K1-48, K1-55, L1-2, L1-6, N1-1858, N1-1860, N1-1861, N1-1864 are provided in SEQ ID NOS: 30-40, respectively. Comparison of these sequences with those in the gene bank as described above, revealed no significant homologies to three of the five most abundant DNA species, (J1-17, L1-12 and N1-1862; SEQ ID NOS: 8-9, 10-11 and 12-13, respectively). Of the remaining two most abundant species, one (J1-12; SEQ ID NO:30) was found to be identical to the previously identified human pulmonary surfactant-associated protein, and the other (K1-48; SEQ ID NO:33) was determined to have some homology to *R. norvegicus* mRNA for 2-arylpropionyl-CoA epimerase. Of the 17 less abundant cDNA clones isolated from the subtracted prostate tumor specific cDNA library with spike, four (J1-16, K1-55, L1-6 and N1-1864; SEQ ID NOS:31, 34, 36 and 40, respectively) were found to be identical to previously identified sequences, two (J1-21 and N1-1860; SEQ ID NOS: 32 and 38, respectively) were found to show some homology to non-human sequences, and two (L1-2 and N1-1861; SEQ ID NOS: 35 and 39, respectively) were found to show some homology to known human sequences. No significant homologies were found to the polypeptides J1-13, J1-19, J1-24, J1-25, K1-58, K1-63, L1-4, L1-14 (SEQ ID NOS: 14-15, 16-17, 20-21, 18-19, 22-23, 24-25, 26-27, 28-29, respectively).

Subsequent studies led to the isolation of full length cDNA sequences for J1-17, L1-12 and N1-1862 (SEQ ID NOS: 109-111, respectively). The corresponding predicted amino acid sequences are provided in SEQ ID NOS: 112-114. L1-12 is also referred to as P501S. A cDNA splice variant of P501S is provided in SEQ ID NO: 606.

In a further experiment, four additional clones were identified by subtracting a prostate tumor cDNA library with normal prostate cDNA prepared from a pool of three normal prostate poly A<sup>+</sup> RNA (referred to as "prostate subtraction 2"). The determined



cDNA sequences for these clones, hereinafter referred to as U1-3064, U1-3065, V1-3692 and 1A-3905, are provided in SEQ ID NO: 69-72, respectively. Comparison of the determined sequences with those in the gene bank revealed no significant homologies to U1-3065.

5                   A second subtraction with spike (referred to as “prostate subtraction spike 2”) was performed by subtracting a prostate tumor specific cDNA library with spike with normal pancreas cDNA library and further spiked with PSA, J1-17, pulmonary surfactant-associated protein, mitochondrial DNA, cytochrome c oxidase subunit II, N1-1862, autonomously replicating sequence, L1-12 and tumor expression enhanced gene. Four  
10 additional clones, hereinafter referred to as V1-3686, R1-2330, 1B-3976 and V1-3679, were isolated. The determined cDNA sequences for these clones are provided in SEQ ID NO:73-76, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to V1-3686 and R1-2330.

                  Further analysis of the three prostate subtractions described above (prostate  
15 subtraction 2, subtracted prostate tumor specific cDNA library with spike, and prostate subtraction spike 2) resulted in the identification of sixteen additional clones, referred to as 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1G-4734, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4810, 1I-4811, 1J-4876, 1K-4884 and 1K-4896. The determined cDNA sequences for these clones are provided in SEQ ID NOS: 77-92, respectively. Comparison  
20 of these sequences with those in the gene bank as described above, revealed no significant homologies to 1G-4741, 1G-4734, 1I-4807, 1J-4876 and 1K-4896 (SEQ ID NOS: 79, 81, 87, 90 and 92, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4807, 1J-4876, 1K-4884 and 1K-  
25 4896, provided in SEQ ID NOS: 179-188 and 191-193, respectively, and to the determination of additional partial cDNA sequences for 1I-4810 and 1I-4811, provided in SEQ ID NOS: 189 and 190, respectively.

                  Additional studies with prostate subtraction spike 2 resulted in the isolation of three more clones. Their sequences were determined as described above and compared



to the most recent GenBank. All three clones were found to have homology to known genes, which are Cysteine-rich protein, KIAA0242, and KIAA0280 (SEQ ID NO: 317, 319, and 320, respectively). Further analysis of these clones by Synteni microarray (Synteni, Palo Alto, CA) demonstrated that all three clones were over-expressed in most prostate tumors and prostate BPH, as well as in the majority of normal prostate tissues tested, but low expression in all other normal tissues.

An additional subtraction was performed by subtracting a normal prostate cDNA library with normal pancreas cDNA (referred to as "prostate subtraction 3"). This led to the identification of six additional clones referred to as 1G-4761, 1G-4762, 1H-4766, 1H-4770, 1H-4771 and 1H-4772 (SEQ ID NOS: 93-98). Comparison of these sequences with those in the gene bank revealed no significant homologies to 1G-4761 and 1H-4771 (SEQ ID NOS: 93 and 97, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4761, 1G-4762, 1H-4766 and 1H-4772 provided in SEQ ID NOS: 194-196 and 199, respectively, and to the determination of additional partial cDNA sequences for 1H-4770 and 1H-4771, provided in SEQ ID NOS: 197 and 198, respectively.

Subtraction of a prostate tumor cDNA library, prepared from a pool of polyA<sup>+</sup> RNA from three prostate cancer patients, with a normal pancreas cDNA library (prostate subtraction 4) led to the identification of eight clones, referred to as 1D-4297, 1D-4309, 1D-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280 (SEQ ID NOS: 99-107). These sequences were compared to those in the gene bank as described above. No significant homologies were found to 1D-4283 and 1D-4304 (SEQ ID NOS: 103 and 104, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1D-4309, 1D-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280, provided in SEQ ID NOS: 200-206, respectively.

cDNA clones isolated in prostate subtraction 1 and prostate subtraction 2, described above, were colony PCR amplified and their mRNA expression levels in prostate tumor, normal prostate and in various other normal tissues were determined using microarray technology (Synteni, Palo Alto, CA). Briefly, the PCR amplification products



were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured.

5 This intensity correlates with the hybridization intensity. Two clones (referred to as P509S and P510S) were found to be over-expressed in prostate tumor and normal prostate and expressed at low levels in all other normal tissues tested (liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon). The determined cDNA sequences for

10 P509S and P510S are provided in SEQ ID NO: 223 and 224, respectively. Comparison of these sequences with those in the gene bank as described above, revealed some homology to previously identified ESTs.

Additional, studies led to the isolation of the full-length cDNA sequence for P509S. This sequence is provided in SEQ ID NO: 332, with the corresponding predicted

15 amino acid sequence being provided in SEQ ID NO: 339. Two variant full-length cDNA sequences for P510S are provided in SEQ ID NO: 535 and 536, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 537 and 538, respectively. Additional splice variants of P510S are provided in SEQ ID NO: 598 and 599.

The determined cDNA sequences for additional prostate-specific clones

20 isolated during characterization of prostate specific cDNA libraries are provided in SEQ ID NO: 618-689, 691-697 and 709-772. Comparison of these sequences with those in the public databases revealed no significant homologies to any of these sequences.

## EXAMPLE 2

### 25 DETERMINATION OF TISSUE SPECIFICITY OF PROSTATE-SPECIFIC POLYPEPTIDES

Using gene specific primers, mRNA expression levels for the representative prostate-specific polypeptides F1-16, H1-1, J1-17 (also referred to as P502S), L1-12 (also



referred to as P501S), F1-12 (also referred to as P504S) and N1-1862 (also referred to as P503S) were examined in a variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent as described above. First strand synthesis was carried out using 1-2  
 5  $\mu$ g of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 °C for one hour. The cDNA was then amplified by PCR with gene-specific primers. To ensure the semi-quantitative nature of the RT-PCR,  $\beta$ -actin was used as an internal control for each of the tissues examined. First, serial dilutions of the first strand cDNAs were prepared and RT-PCR assays were performed using  $\beta$ -actin specific primers. A dilution  
 10 was then chosen that enabled the linear range amplification of the  $\beta$ -actin template and which was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the  $\beta$ -actin levels were determined for each reverse transcription reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding  
 15 reverse transcriptase.

mRNA Expression levels were examined in four different types of tumor tissue (prostate tumor from 2 patients, breast tumor from 3 patients, colon tumor, lung tumor), and sixteen different normal tissues, including prostate, colon, kidney, liver, lung, ovary, pancreas, skeletal muscle, skin, stomach, testes, bone marrow and brain. F1-16 was  
 20 found to be expressed at high levels in prostate tumor tissue, colon tumor and normal prostate, and at lower levels in normal liver, skin and testes, with expression being undetectable in the other tissues examined. H1-1 was found to be expressed at high levels in prostate tumor, lung tumor, breast tumor, normal prostate, normal colon and normal brain, at much lower levels in normal lung, pancreas, skeletal muscle, skin, small intestine,  
 25 bone marrow, and was not detected in the other tissues tested. J1-17 (P502S) and L1-12 (P501S) appear to be specifically over-expressed in prostate, with both genes being expressed at high levels in prostate tumor and normal prostate but at low to undetectable levels in all the other tissues examined. N1-1862 (P503S) was found to be over-expressed in 60% of prostate tumors and detectable in normal colon and kidney. The RT-PCR results



thus indicate that F1-16, H1-1, J1-17 (P502S), N1-1862 (P503S) and L1-12 (P501S) are either prostate specific or are expressed at significantly elevated levels in prostate.

Further RT-PCR studies showed that F1-12 (P504S) is over-expressed in 60% of prostate tumors, detectable in normal kidney but not detectable in all other tissues tested. Similarly, R1-2330 was shown to be over-expressed in 40% of prostate tumors, detectable in normal kidney and liver, but not detectable in all other tissues tested. U1-3064 was found to be over-expressed in 60% of prostate tumors, and also expressed in breast and colon tumors, but was not detectable in normal tissues.

RT-PCR characterization of R1-2330, U1-3064 and 1D-4279 showed that these three antigens are over-expressed in prostate and/or prostate tumors.

Northern analysis with four prostate tumors, two normal prostate samples, two BPH prostates, and normal colon, kidney, liver, lung, pancreas, skeletal muscle, brain, stomach, testes, small intestine and bone marrow, showed that L1-12 (P501S) is over-expressed in prostate tumors and normal prostate, while being undetectable in other normal tissues tested. J1-17 (P502S) was detected in two prostate tumors and not in the other tissues tested. N1-1862 (P503S) was found to be over-expressed in three prostate tumors and to be expressed in normal prostate, colon and kidney, but not in other tissues tested. F1-12 (P504S) was found to be highly expressed in two prostate tumors and to be undetectable in all other tissues tested.

The microarray technology described above was used to determine the expression levels of representative antigens described herein in prostate tumor, breast tumor and the following normal tissues: prostate, liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon. L1-12 (P501S) was found to be over-expressed in normal prostate and prostate tumor, with some expression being detected in normal skeletal muscle. Both J1-12 and F1-12 (P504S) were found to be over-expressed in prostate tumor, with expression being lower or undetectable in all other tissues tested. N1-1862 (P503S) was found to be expressed at high levels in prostate tumor and normal prostate, and at low levels in normal large intestine and normal colon, with expression



being undetectable in all other tissues tested. R1-2330 was found to be over-expressed in prostate tumor and normal prostate, and to be expressed at lower levels in all other tissues tested. 1D-4279 was found to be over-expressed in prostate tumor and normal prostate, expressed at lower levels in normal spinal cord, and to be undetectable in all other tissues  
5 tested.

Further microarray analysis to specifically address the extent to which P501S (SEQ ID NO: 110) was expressed in breast tumor revealed moderate over-expression not only in breast tumor, but also in metastatic breast tumor (2/31), with negligible to low expression in normal tissues. This data suggests that P501S may be over-  
10 expressed in various breast tumors as well as in prostate tumors.

The expression levels of 32 ESTs (expressed sequence tags) described by Vasmatazis *et al.* (*Proc. Natl. Acad. Sci. USA* 95:300-304, 1998) in a variety of tumor and normal tissues were examined by microarray technology as described above. Two of these clones (referred to as P1000C and P1001C) were found to be over-expressed in prostate  
15 tumor and normal prostate, and expressed at low to undetectable levels in all other tissues tested (normal aorta, thymus, resting and activated PBMC, epithelial cells, spinal cord, adrenal gland, fetal tissues, skin, salivary gland, large intestine, bone marrow, liver, lung, dendritic cells, stomach, lymph nodes, brain, heart, small intestine, skeletal muscle, colon and kidney. The determined cDNA sequences for P1000C and P1001C are provided in  
20 SEQ ID NO: 384 and 472, respectively. The sequence of P1001C was found to show some homology to the previously isolated Human mRNA for JM27 protein. No significant homologies were found to the sequence of P1000C.

The expression of the polypeptide encoded by the full length cDNA sequence for F1-12 (also referred to as P504S; SEQ ID NO: 108) was investigated by  
25 immunohistochemical analysis. Rabbit-anti-P504S polyclonal antibodies were generated against the full length P504S protein by standard techniques. Subsequent isolation and characterization of the polyclonal antibodies were also performed by techniques well known in the art. Immunohistochemical analysis showed that the P504S polypeptide was expressed in 100% of prostate carcinoma samples tested (n=5).



The rabbit-anti-P504S polyclonal antibody did not appear to label benign prostate cells with the same cytoplasmic granular staining, but rather with light nuclear staining. Analysis of normal tissues revealed that the encoded polypeptide was found to be expressed in some, but not all normal human tissues. Positive cytoplasmic staining with  
 5 rabbit-anti-P504S polyclonal antibody was found in normal human kidney, liver, brain, colon and lung-associated macrophages, whereas heart and bone marrow were negative.

This data indicates that the P504S polypeptide is present in prostate cancer tissues, and that there are qualitative and quantitative differences in the staining between benign prostatic hyperplasia tissues and prostate cancer tissues, suggesting that this  
 10 polypeptide may be detected selectively in prostate tumors and therefore be useful in the diagnosis of prostate cancer.

### EXAMPLE 3

#### ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC 15 POLYPEPTIDES BY PCR-BASED SUBTRACTION

A cDNA subtraction library, containing cDNA from normal prostate subtracted with ten other normal tissue cDNAs (brain, heart, kidney, liver, lung, ovary, placenta, skeletal muscle, spleen and thymus) and then submitted to a first round of PCR  
 20 amplification, was purchased from Clontech. This library was subjected to a second round of PCR amplification, following the manufacturer's protocol. The resulting cDNA fragments were subcloned into the vector pT7 Blue T-vector (Novagen, Madison, WI) and transformed into XL-1 Blue MRF' *E. coli* (Stratagene). DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division  
 25 Automated Sequencer Model 373A.

Fifty-nine positive clones were sequenced. Comparison of the DNA sequences of these clones with those in the gene bank, as described above, revealed no significant homologies to 25 of these clones, hereinafter referred to as P5, P8, P9, P18, P20, P30, P34, P36, P38, P39, P42, P49, P50, P53, P55, P60, P64, P65, P73, P75, P76, P79



and P84. The determined cDNA sequences for these clones are provided in SEQ ID NO: 41-45, 47-52 and 54-65, respectively. P29, P47, P68, P80 and P82 (SEQ ID NO: 46, 53 and 66-68, respectively) were found to show some degree of homology to previously identified DNA sequences. To the best of the inventors' knowledge, none of these  
5 sequences have been previously shown to be present in prostate.

Further studies employing the sequence of SEQ ID NO: 67 as a probe in standard full-length cloning methods, resulted in the isolation of three cDNA sequences which appear to be splice variants of P80 (also known as P704P). These sequences are provided in SEQ ID NO: 699-701.

10 Further studies using the PCR-based methodology described above resulted in the isolation of more than 180 additional clones, of which 23 clones were found to show no significant homologies to known sequences. The determined cDNA sequences for these clones are provided in SEQ ID NO: 115-123, 127, 131, 137, 145, 147-151, 153, 156-158 and 160. Twenty-three clones (SEQ ID NO: 124-126, 128-130, 132-136, 138-144, 146,  
15 152, 154, 155 and 159) were found to show some homology to previously identified ESTs. An additional ten clones (SEQ ID NO: 161-170) were found to have some degree of homology to known genes. Larger cDNA clones containing the P20 sequence represent splice variants of a gene referred to as P703P. The determined DNA sequence for the variants referred to as DE1, DE13 and DE14 are provided in SEQ ID NOS: 171, 175 and  
20 177, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 172, 176 and 178, respectively. The determined cDNA sequence for an extended spliced form of P703 is provided in SEQ ID NO: 225. The DNA sequences for the splice variants referred to as DE2 and DE6 are provided in SEQ ID NOS: 173 and 174, respectively.

25 mRNA Expression levels for representative clones in tumor tissues (prostate (n=5), breast (n=2), colon and lung) normal tissues (prostate (n=5), colon, kidney, liver, lung (n=2), ovary (n=2), skeletal muscle, skin, stomach, small intestine and brain), and activated and non-activated PBMC was determined by RT-PCR as described above. Expression was examined in one sample of each tissue type unless otherwise indicated.



P9 was found to be highly expressed in normal prostate and prostate tumor compared to all normal tissues tested except for normal colon which showed comparable expression. P20, a portion of the P703P gene, was found to be highly expressed in normal prostate and prostate tumor, compared to all twelve normal tissues tested. A modest increase in expression of P20 in breast tumor (n=2), colon tumor and lung tumor was seen compared to all normal tissues except lung (1 of 2). Increased expression of P18 was found in normal prostate, prostate tumor and breast tumor compared to other normal tissues except lung and stomach. A modest increase in expression of P5 was observed in normal prostate compared to most other normal tissues. However, some elevated expression was seen in normal lung and PBMC. Elevated expression of P5 was also observed in prostate tumors (2 of 5), breast tumor and one lung tumor sample. For P30, similar expression levels were seen in normal prostate and prostate tumor, compared to six of twelve other normal tissues tested. Increased expression was seen in breast tumors, one lung tumor sample and one colon tumor sample, and also in normal PBMC. P29 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to the majority of normal tissues. However, substantial expression of P29 was observed in normal colon and normal lung (2 of 2). P80 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to all other normal tissues tested, with increased expression also being seen in colon tumor.

Further studies resulted in the isolation of twelve additional clones, hereinafter referred to as 10-d8, 10-h10, 11-c8, 7-g6, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3, 8-h11, 9-f12 and 9-f3. The determined DNA sequences for 10-d8, 10-h10, 11-c8, 8-d4, 8-d9, 8-h11, 9-f12 and 9-f3 are provided in SEQ ID NO: 207, 208, 209, 216, 217, 220, 221 and 222, respectively. The determined forward and reverse DNA sequences for 7-g6, 8-b5, 8-b6 and 8-g3 are provided in SEQ ID NO: 210 and 211; 212 and 213; 214 and 215; and 218 and 219, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to the sequence of 9-f3. The clones 10-d8, 11-c8 and 8-h11 were found to show some homology to previously isolated ESTs, while 10-h10, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3 and 9-f12 were found to show some homology to previously identified genes.



Further characterization of 7-G6 and 8-G3 showed identity to the known genes PAP and PSA, respectively.

mRNA expression levels for these clones were determined using the micro-array technology described above. The clones 7-G6, 8-G3, 8-B5, 8-B6, 8-D4, 8-D9, 9-F3, 9-F12, 9-H3, 10-A2, 10-A4, 11-C9 and 11-F2 were found to be over-expressed in prostate tumor and normal prostate, with expression in other tissues tested being low or undetectable. Increased expression of 8-F11 was seen in prostate tumor and normal prostate, bladder, skeletal muscle and colon. Increased expression of 10-H10 was seen in prostate tumor and normal prostate, bladder, lung, colon, brain and large intestine. Increased expression of 9-B1 was seen in prostate tumor, breast tumor, and normal prostate, salivary gland, large intestine and skin, with increased expression of 11-C8 being seen in prostate tumor, and normal prostate and large intestine.

An additional cDNA fragment derived from the PCR-based normal prostate subtraction, described above, was found to be prostate specific by both micro-array technology and RT-PCR. The determined cDNA sequence of this clone (referred to as 9-A11) is provided in SEQ ID NO: 226. Comparison of this sequence with those in the public databases revealed 99% identity to the known gene HOXB13.

Further studies led to the isolation of the clones 8-C6 and 8-H7. The determined cDNA sequences for these clones are provided in SEQ ID NO: 227 and 228, respectively. These sequences were found to show some homology to previously isolated ESTs.

PCR and hybridization-based methodologies were employed to obtain longer cDNA sequences for clone P20 (also referred to as P703P), yielding three additional cDNA fragments that progressively extend the 5' end of the gene. These fragments, referred to as P703PDE5, P703P6.26, and P703PX-23 (SEQ ID NO: 326, 328 and 330, with the predicted corresponding amino acid sequences being provided in SEQ ID NO: 327, 329 and 331, respectively) contain additional 5' sequence. P703PDE5 was recovered by screening of a cDNA library (#141-26) with a portion of P703P as a probe. P703P6.26 was recovered from a mixture of three prostate tumor cDNAs and P703PX\_23 was



recovered from cDNA library (#438-48). Together, the additional sequences include all of the putative mature serine protease along with part of the putative signal sequence. The full-length cDNA sequence for P703P is provided in SEQ ID NO: 524, with the corresponding amino acid sequence being provided in SEQ ID NO: 525.

5 P703P was found to show some homology to previously identified proteases, such as thrombin. The thrombin receptor has been shown to be preferentially expressed in highly metastatic breast carcinoma cells and breast carcinoma biopsy samples. Introduction of thrombin receptor antisense cDNA has been shown to inhibit the invasion of metastatic breast carcinoma cells in culture. Antibodies against thrombin receptor  
10 inhibit thrombin receptor activation and thrombin-induced platelet activation. Furthermore, peptides that resemble the receptor's tethered ligand domain inhibit platelet aggregation by thrombin. P703P may play a role in prostate cancer through a protease-activated receptor on the cancer cell or on stromal cells. The potential trypsin-like protease activity of P703P may either activate a protease-activated receptor on the cancer cell  
15 membrane to promote tumorigenesis or activate a protease-activated receptor on the adjacent cells (such as stromal cells) to secrete growth factors and/or proteases (such as matrix metalloproteinases) that could promote tumor angiogenesis, invasion and metastasis. P703P may thus promote tumor progression and/or metastasis through the activation of protease-activated receptor. Polypeptides and antibodies that block the  
20 P703P-receptor interaction may therefore be usefully employed in the treatment of prostate cancer.

Further studies using a PCR-based subtraction library of a prostate tumor pool subtracted against a pool of normal tissues (referred to as JP: PCR subtraction) resulted in the isolation of thirteen additional clones, seven of which did not share any  
25 significant homology to known GenBank sequences. The determined cDNA sequences for these seven clones (P711P, P712P, novel 23, P774P, P775P, P710P and P768P) are provided in SEQ ID NO: 307-311, 313 and 315, respectively. The remaining six clones (SEQ ID NO: 316 and 321-325) were shown to share some homology to known genes. By microarray analysis, all thirteen clones showed three or more fold over-expression in



prostate tissues, including prostate tumors, BPH and normal prostate as compared to normal non-prostate tissues. Clones P711P, P712P, novel 23 and P768P showed over-expression in most prostate tumors and BPH tissues tested (n=29), and in the majority of normal prostate tissues (n=4), but background to low expression levels in all normal  
 5 tissues. Clones P774P, P775P and P710P showed comparatively lower expression and expression in fewer prostate tumors and BPH samples, with negative to low expression in normal prostate.

Further studies led to the isolation of an extended cDNA sequence for P712P (SEQ ID NO: 552). The amino acid sequences encoded by 16 predicted open  
 10 reading frames present within the sequence of SEQ ID NO: 552 are provided in SEQ ID NO: 553-568.

The full-length cDNA for P711P was obtained by employing the partial sequence of SEQ ID NO: 307 to screen a prostate cDNA library. Specifically, a directionally cloned prostate cDNA library was prepared using standard techniques. One  
 15 million colonies of this library were plated onto LB/Amp plates. Nylon membrane filters were used to lift these colonies, and the cDNAs which were picked up by these filters were denatured and cross-linked to the filters by UV light. The P711P cDNA fragment of SEQ ID NO: 307 was radio-labeled and used to hybridize with these filters. Positive clones were selected, and cDNAs were prepared and sequenced using an automatic Perkin  
 20 Elmer/Applied Biosystems sequencer. The determined full-length sequence of P711P is provided in SEQ ID NO: 382, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 383.

Using PCR and hybridization-based methodologies, additional cDNA sequence information was derived for two clones described above, 11-C9 and 9-F3, herein  
 25 after referred to as P707P and P714P, respectively (SEQ ID NO: 333 and 334). After comparison with the most recent GenBank, P707P was found to be a splice variant of the known gene HoxB13. In contrast, no significant homologies to P714P were found. Further studies employing the sequence of SEQ ID NO: 334 as a probe in standard full-length cloning methods, resulted in an extended cDNA sequence for P714P. This sequence



is provided in SEQ ID NO: 698. This sequence was found to show some homology to the gene that encodes human ribosomal L23A protein.

Clones 8-B3, P89, P98, P130 and P201 (as disclosed in U.S. Patent Application No. 09/020,956, filed February 9, 1998) were found to be contained within one  
 5 contiguous sequence, referred to as P705P (SEQ ID NO: 335, with the predicted amino acid sequence provided in SEQ ID NO: 336), which was determined to be a splice variant of the known gene NKX 3.1.

Further studies on P775P resulted in the isolation of four additional sequences (SEQ ID NO: 473-476) which are all splice variants of the P775P gene. The  
 10 sequence of SEQ ID NO: 474 was found to contain two open reading frames (ORFs). The predicted amino acid sequences encoded by these ORFs are provided in SEQ ID NO: 477 and 478. The cDNA sequence of SEQ ID NO: 475 was found to contain an ORF which encodes the amino acid sequence of SEQ ID NO: 479. The cDNA sequence of SEQ ID NO: 473 was found to contain four ORFs. The predicted amino acid sequences encoded by  
 15 these ORFs are provided in SEQ ID NO: 480-483. Additional splice variants of P775P are provided in SEQ ID NO: 593-597.

Subsequent studies led to the identification of a genomic region on chromosome 22q11.2, known as the Cat Eye Syndrome region, that contains the five prostate genes P704P, P712P, P774P, P775P and B305D. The relative location of each of  
 20 these five genes within the genomic region is shown in Fig. 10. This region may therefore be associated with malignant tumors, and other potential tumor genes may be contained within this region. These studies also led to the identification of a potential open reading frame (ORF) for P775P (provided in SEQ ID NO: 533), which encodes the amino acid sequence of SEQ ID NO: 534.

25 Comparison of the clone of SEQ ID NO: 325 (referred to as P558S) with sequences in the GenBank and GeneSeq DNA databases showed that P558S is identical to the prostate-specific transglutaminase gene, which is known to have two forms. The full-length sequences for the two forms are provided in SEQ ID NO: 773 and 774, with the corresponding amino acid sequences being provided in SEQ ID NO: 775 and 776,



respectively. The cDNA sequence of SEQ ID NO: 774 has a 15 pair base insert, resulting in a 5 amino acid insert in the corresponding amino acid sequence (SEQ ID NO: 776). This insert is not present in the sequence of SEQ ID NO: 773.

5

## EXAMPLE 4

## SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems 430A peptide synthesizer using Fmoc chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

15

20

## EXAMPLE 5

FURTHER ISOLATION AND CHARACTERIZATION OF  
PROSTATE-SPECIFIC POLYPEPTIDES BY PCR-BASED SUBTRACTION

25

A cDNA library generated from prostate primary tumor mRNA as described above was subtracted with cDNA from normal prostate. The subtraction was performed using a PCR-based protocol (Clontech), which was modified to generate larger fragments. Within this protocol, tester and driver double stranded cDNA were separately digested with



five restriction enzymes that recognize six-nucleotide restriction sites (MluI, MscI, PvuII, SalI and StuI). This digestion resulted in an average cDNA size of 600 bp, rather than the average size of 300 bp that results from digestion with RsaI according to the Clontech protocol. This modification did not affect the subtraction efficiency. Two tester  
 5 populations were then created with different adapters, and the driver library remained without adapters.

The tester and driver libraries were then hybridized using excess driver cDNA. In the first hybridization step, driver was separately hybridized with each of the two tester cDNA populations. This resulted in populations of (a) unhybridized tester  
 10 cDNAs, (b) tester cDNAs hybridized to other tester cDNAs, (c) tester cDNAs hybridized to driver cDNAs and (d) unhybridized driver cDNAs. The two separate hybridization reactions were then combined, and rehybridized in the presence of additional denatured driver cDNA. Following this second hybridization, in addition to populations (a) through (d), a fifth population (e) was generated in which tester cDNA with one adapter hybridized  
 15 to tester cDNA with the second adapter. Accordingly, the second hybridization step resulted in enrichment of differentially expressed sequences which could be used as templates for PCR amplification with adaptor-specific primers.

The ends were then filled in, and PCR amplification was performed using adaptor-specific primers. Only population (e), which contained tester cDNA that did not  
 20 hybridize to driver cDNA, was amplified exponentially. A second PCR amplification step was then performed, to reduce background and further enrich differentially expressed sequences.

This PCR-based subtraction technique normalizes differentially expressed cDNAs so that rare transcripts that are overexpressed in prostate tumor tissue may be  
 25 recoverable. Such transcripts would be difficult to recover by traditional subtraction methods.

In addition to genes known to be overexpressed in prostate tumor, seventy-seven further clones were identified. Sequences of these partial cDNAs are provided in SEQ ID NO: 29 to 305. Most of these clones had no significant homology to database



sequences. Exceptions were JTPN23 (SEQ ID NO: 231; similarity to pig valosin-containing protein), JTPN30 (SEQ ID NO: 234; similarity to rat mRNA for proteasome subunit), JTPN45 (SEQ ID NO: 243; similarity to rat *norvegicus* cytosolic NADP-dependent isocitrate dehydrogenase), JTPN46 (SEQ ID NO: 244; similarity to human subclone H8 4 d4 DNA sequence), JP1D6 (SEQ ID NO: 265; similarity to *G. gallus* dynein light chain-A), JP8D6 (SEQ ID NO: 288; similarity to human BAC clone RG016J04), JP8F5 (SEQ ID NO: 289; similarity to human subclone H8 3 b5 DNA sequence), and JP8E9 (SEQ ID NO: 299; similarity to human Alu sequence).

Additional studies using the PCR-based subtraction library consisting of a prostate tumor pool subtracted against a normal prostate pool (referred to as PT-PN PCR subtraction) yielded three additional clones. Comparison of the cDNA sequences of these clones with the most recent release of GenBank revealed no significant homologies to the two clones referred to as P715P and P767P (SEQ ID NO: 312 and 314). The remaining clone was found to show some homology to the known gene KIAA0056 (SEQ ID NO: 318). Using microarray analysis to measure mRNA expression levels in various tissues, all three clones were found to be over-expressed in prostate tumors and BPH tissues. Specifically, clone P715P was over-expressed in most prostate tumors and BPH tissues by a factor of three or greater, with elevated expression seen in the majority of normal prostate samples and in fetal tissue, but negative to low expression in all other normal tissues. Clone P767P was over-expressed in several prostate tumors and BPH tissues, with moderate expression levels in half of the normal prostate samples, and background to low expression in all other normal tissues tested.

Further analysis, by microarray as described above, of the PT-PN PCR subtraction library and of a DNA subtraction library containing cDNA from prostate tumor subtracted with a pool of normal tissue cDNAs, led to the isolation of 27 additional clones (SEQ ID NO: 340-365 and 381) which were determined to be over-expressed in prostate tumor. The clones of SEQ ID NO: 341, 342, 345, 347, 348, 349, 351, 355-359, 361, 362 and 364 were also found to be expressed in normal prostate. Expression of all 26 clones in a variety of normal tissues was found to be low or undetectable, with the exception of



P544S (SEQ ID NO: 356) which was found to be expressed in small intestine. Of the 26 clones, 11 (SEQ ID NO: 340-349 and 362) were found to show some homology to previously identified sequences. No significant homologies were found to the clones of SEQ ID NO: 350, 351, 353-361, and 363-365.

5                   Comparison of the sequence of SEQ ID NO: 362 with sequences in the GenBank and GeneSeq DNA databases showed that this clone (referred to as P788P) is identical to GeneSeq Accession No. X27262, which encodes a protein found in the GeneSeq protein Accession No. Y00931. The full length cDNA sequence of P788P is shown in Figure 12A (SEQ ID NO: 777), with the corresponding predicted amino acid  
10   being shown in Figure 12B (SEQ ID NO: 778). Subsequently, a full-length cDNA sequence for P788P that contains polymorphisms not found in the sequence of SEQ ID NO: 779, was cloned multiple times by PCR amplification from cDNA prepared from several RNA templates from three individuals. This determined cDNA sequence of this polymorphic variant of P788P is provided in SEQ ID NO: 779, with the corresponding  
15   amino acid sequence being provided in SEQ ID NO: 780. The sequence of SEQ ID NO: 780 differs from that of SEQ ID NO: 778 by six amino acid residues.

                  Further studies on the clone of SEQ ID NO: 352 (referred to as P790P) led to the isolation of the full-length cDNA sequence of SEQ ID NO: 526. The corresponding predicted amino acid is provided in SEQ ID NO: 527. Data from two quantitative PCR  
20   experiments indicated that P790P is over-expressed in 11/15 tested prostate tumor samples and is expressed at low levels in spinal cord, with no expression being seen in all other normal samples tested. Data from further PCR experiments and microarray experiments showed over-expression in normal prostate and prostate tumor with little or no expression in other tissues tested. P790P was subsequently found to show significant homology to a  
25   previously identified G-protein coupled prostate tissue receptor.

                  Additional studies on the clone of SEQ ID NO: 354 (referred to as P776P) led to the isolation of an extended cDNA sequence, provided in SEQ ID NO: 569. The determined cDNA sequences of three additional splice variants of P776P are provided in SEQ ID NO: 570-572. The amino acid sequences encoded by two predicted open reading



frames (ORFs) contained within SEQ ID NO: 570, one predicted ORF contained within SEQ ID NO: 571, and 11 predicted ORFs contained within SEQ ID NO: 569, are provided in SEQ ID NO: 573-586, respectively.

Comparison of the cDNA sequences for the clones P767P (SEQ ID NO: 314) and P777P (SEQ ID NO: 350) with sequences in the GenBank human EST database showed that the two clones matched many EST sequences in common, suggesting that P767P and P777P may represent the same gene. A DNA consensus sequence derived from a DNA sequence alignment of P767P, P777P and multiple EST clones is provided in SEQ ID NO: 587. The amino acid sequences encoded by three putative ORFs located within SEQ ID NO: 587 are provided in SEQ ID NO: 588-590.

## EXAMPLE 6

### PEPTIDE PRIMING OF MICE AND PROPAGATION OF CTL LINES

6.1. This Example illustrates the preparation of a CTL cell line specific for cells expressing the P502S gene.

Mice expressing the transgene for human HLA A2Kb (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with P2S#12 peptide (VLGWVAEL; SEQ ID NO: 306), which is derived from the P502S gene (also referred to herein as J1-17, SEQ ID NO: 8), as described by Theobald et al., *Proc. Natl. Acad. Sci. USA* 92:11993-11997, 1995 with the following modifications. Mice were immunized with 100µg of P2S#12 and 120µg of an I-A<sup>b</sup> binding peptide derived from hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and using a nylon mesh single cell suspensions prepared. Cells were then resuspended at  $6 \times 10^6$  cells/ml in complete media (RPMI-1640; Gibco BRL, Gaithersburg, MD) containing 10% FCS, 2mM Glutamine (Gibco BRL), sodium pyruvate (Gibco BRL), non-essential amino acids (Gibco BRL),  $2 \times 10^{-5}$  M 2-mercaptoethanol, 50U/ml penicillin and streptomycin, and cultured in the presence of irradiated (3000 rads) P2S#12-pulsed (5mg/ml P2S#12 and 10mg/ml β2-microglobulin) LPS blasts (A2



transgenic spleens cells cultured in the presence of 7µg/ml dextran sulfate and 25µg/ml LPS for 3 days). Six days later, cells ( $5 \times 10^5$ /ml) were restimulated with  $2.5 \times 10^6$ /ml peptide pulsed irradiated (20,000 rads) EL4A2Kb cells (Sherman et al, *Science* 258:815-818, 1992) and  $3 \times 10^6$ /ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20U/ml IL-2. Cells continued to be restimulated on a weekly basis as described, in preparation for cloning the line.

P2S#12 line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells ( $1 \times 10^4$  cells/ well) as stimulators and A2 transgenic spleen cells as feeders ( $5 \times 10^5$  cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, clones that were growing were isolated and maintained in culture. Several of these clones demonstrated significantly higher reactivity (lysis) against human fibroblasts (HLA A2Kb expressing) transduced with P502S than against control fibroblasts. An example is presented in Figure 1.

This data indicates that P2S #12 represents a naturally processed epitope of the P502S protein that is expressed in the context of the human HLA A2Kb molecule.

6.2. This Example illustrates the preparation of murine CTL lines and CTL clones specific for cells expressing the P501S gene.

This series of experiments were performed similarly to that described above. Mice were immunized with the P1S#10 peptide (SEQ ID NO: 337), which is derived from the P501S gene (also referred to herein as L1-12, SEQ ID NO: 110). The P1S#10 peptide was derived by analysis of the predicted polypeptide sequence for P501S for potential HLA-A2 binding sequences as defined by published HLA-A2 binding motifs (Parker, KC, et al, *J. Immunol.*, 152:163, 1994). P1S#10 peptide was synthesized as described in Example 4, and empirically tested for HLA-A2 binding using a T cell based competition assay. Predicted A2 binding peptides were tested for their ability to compete HLA-A2 specific peptide presentation to an HLA-A2 restricted CTL clone (D150M58), which is specific for the HLA-A2 binding influenza matrix peptide fluM58. D150M58 CTL secretes TNF in response to self-presentation of peptide fluM58. In the competition assay,



test peptides at 100-200 µg/ml were added to cultures of D150M58 CTL in order to bind HLA-A2 on the CTL. After thirty minutes, CTL cultured with test peptides, or control peptides, were tested for their antigen dose response to the fluM58 peptide in a standard TNF bioassay. As shown in Figure 3, peptide P1S#10 competes HLA-A2 restricted presentation of fluM58, demonstrating that peptide P1S#10 binds HLA-A2.

Mice expressing the transgene for human HLA A2Kb were immunized as described by Theobald et al. (*Proc. Natl. Acad. Sci. USA* 92:11993-11997, 1995) with the following modifications. Mice were immunized with 62.5µg of P1S #10 and 120µg of an I-A<sup>b</sup> binding peptide derived from Hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and single cell suspensions prepared using a nylon mesh. Cells were then resuspended at  $6 \times 10^6$  cells/ml in complete media (as described above) and cultured in the presence of irradiated (3000 rads) P1S#10-pulsed (2µg/ml P1S#10 and 10mg/ml β2-microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7µg/ml dextran sulfate and 25µg/ml LPS for 3 days). Six days later cells ( $5 \times 10^5$ /ml) were restimulated with  $2.5 \times 10^6$ /ml peptide-pulsed irradiated (20,000 rads) EL4A2Kb cells, as described above, and  $3 \times 10^6$ /ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20 U/ml IL-2. Cells were restimulated on a weekly basis in preparation for cloning. After three rounds of *in vitro* stimulations, one line was generated that recognized P1S#10-pulsed Jurkat A2Kb targets and P501S-transduced Jurkat targets as shown in Figure 4.

A P1S#10-specific CTL line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells ( $1 \times 10^4$  cells/ well) as stimulators and A2 transgenic spleen cells as feeders ( $5 \times 10^5$  cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, viable clones were isolated and maintained in culture. As shown in Figure 5, five of these clones demonstrated specific cytolytic reactivity against P501S-transduced Jurkat A2Kb targets. This data indicates that P1S#10 represents a naturally processed epitope of the P501S protein that is expressed in the context of the human HLA-A2.1 molecule.



## EXAMPLE 7

PRIMING OF CTL *IN VIVO* USING NAKED DNA IMMUNIZATION

## WITH A PROSTATE ANTIGEN

The prostate-specific antigen L1-12, as described above, is also referred to  
 5 as P501S. HLA A2Kb Tg mice (provided by Dr L. Sherman, The Scripps Research  
 Institute, La Jolla, CA) were immunized with 100 µg P501S in the vector VR1012 either  
 intramuscularly or intradermally. The mice were immunized three times, with a two week  
 interval between immunizations. Two weeks after the last immunization, immune spleen  
 cells were cultured with Jurkat A2Kb-P501S transduced stimulator cells. CTL lines were  
 10 stimulated weekly. After two weeks of *in vitro* stimulation, CTL activity was assessed  
 against P501S transduced targets. Two out of 8 mice developed strong anti-P501S CTL  
 responses. These results demonstrate that P501S contains at least one naturally processed  
 HLA-A2-restricted CTL epitope.

15

## EXAMPLE 8

## ABILITY OF HUMAN T CELLS TO RECOGNIZE PROSTATE-SPECIFIC POLYPEPTIDES

This Example illustrates the ability of T cells specific for a prostate tumor  
 polypeptide to recognize human tumor.

20

Human CD8<sup>+</sup> T cells were primed *in vitro* to the P2S-12 peptide (SEQ ID  
 NO: 306) derived from P502S (also referred to as J1-17) using dendritic cells according to  
 the protocol of Van Tsai et al. (*Critical Reviews in Immunology* 18:65-75, 1998). The  
 resulting CD8<sup>+</sup> T cell microcultures were tested for their ability to recognize the P2S-12  
 peptide presented by autologous fibroblasts or fibroblasts which were transduced to express  
 25 the P502S gene in a  $\gamma$ -interferon ELISPOT assay (see Lalvani et al., *J. Exp. Med.* 186:859-  
 865, 1997). Briefly, titrating numbers of T cells were assayed in duplicate on 10<sup>4</sup>  
 fibroblasts in the presence of 3 µg/ml human  $\beta_2$ -microglobulin and 1 µg/ml P2S-12 peptide  
 or control E75 peptide. In addition, T cells were simultaneously assayed on autologous  
 fibroblasts transduced with the P502S gene or as a control, fibroblasts transduced with



HER-2/*neu*. Prior to the assay, the fibroblasts were treated with 10 ng/ml  $\gamma$ -interferon for 48 hours to upregulate class I MHC expression. One of the microcultures (#5) demonstrated strong recognition of both peptide pulsed fibroblasts as well as transduced fibroblasts in a  $\gamma$ -interferon ELISPOT assay. Figure 2A demonstrates that there was a strong increase in the number of  $\gamma$ -interferon spots with increasing numbers of T cells on fibroblasts pulsed with the P2S-12 peptide (solid bars) but not with the control E75 peptide (open bars). This shows the ability of these T cells to specifically recognize the P2S-12 peptide. As shown in Figure 2B, this microculture also demonstrated an increase in the number of  $\gamma$ -interferon spots with increasing numbers of T cells on fibroblasts transduced to express the P502S gene but not the HER-2/*neu* gene. These results provide additional confirmatory evidence that the P2S-12 peptide is a naturally processed epitope of the P502S protein. Furthermore, this also demonstrates that there exists in the human T cell repertoire, high affinity T cells which are capable of recognizing this epitope. These T cells should also be capable of recognizing human tumors which express the P502S gene.

## EXAMPLE 9

### ELICITATION OF PROSTATE ANTIGEN-SPECIFIC CTL RESPONSES IN HUMAN BLOOD

This Example illustrates the ability of a prostate-specific antigen to elicit a CTL response in blood of normal humans.

Autologous dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal donors by growth for five days in RPMI medium containing 10% human serum, 50 ng/ml GMCSF and 30 ng/ml IL-4. Following culture, DC were infected overnight with recombinant P501S-expressing vaccinia virus at an M.O.I. of 5 and matured for 8 hours by the addition of 2 micrograms/ml CD40 ligand. Virus was inactivated by UV irradiation, CD8<sup>+</sup> cells were isolated by positive selection using magnetic beads, and priming cultures were initiated in 24-well plates. Following five stimulation cycles using autologous fibroblasts retrovirally transduced to express P501S



and CD80, CD8+ lines were identified that specifically produced interferon-gamma when stimulated with autologous P501S-transduced fibroblasts. The P501S-specific activity of cell line 3A-1 could be maintained following additional stimulation cycles on autologous B-LCL transduced with P501S. Line 3A-1 was shown to specifically recognize autologous B-LCL transduced to express P501S, but not EGFP-transduced autologous B-LCL, as measured by cytotoxicity assays ( $^{51}\text{Cr}$  release) and interferon-gamma production (Interferon-gamma Elispot; *see above and Lalvani et al., J. Exp. Med. 186:859-865, 1997*). The results of these assays are presented in Figures 6A and 6B.

10

## EXAMPLE 10

IDENTIFICATION OF A NATURALLY PROCESSED CTL EPITOPE CONTAINED WITHIN A  
PROSTATE-SPECIFIC ANTIGEN

The 9-mer peptide p5 (SEQ ID NO: 338) was derived from the P703P antigen (also referred to as P20). The p5 peptide is immunogenic in human HLA-A2 donors and is a naturally processed epitope. Antigen specific human CD8+ T cells can be primed following repeated *in vitro* stimulations with monocytes pulsed with p5 peptide. These CTL specifically recognize p5-pulsed and P703P-transduced target cells in both ELISPOT (as described above) and chromium release assays. Additionally, immunization of HLA-A2Kb transgenic mice with p5 leads to the generation of CTL lines which recognize a variety of HLA-A2Kb or HLA-A2 transduced target cells expressing P703P.

Initial studies demonstrating that p5 is a naturally processed epitope were done using HLA-A2Kb transgenic mice. HLA-A2Kb transgenic mice were immunized subcutaneously in the footpad with 100  $\mu\text{g}$  of p5 peptide together with 140  $\mu\text{g}$  of hepatitis B virus core peptide (a Th peptide) in Freund's incomplete adjuvant. Three weeks post immunization, spleen cells from immunized mice were stimulated *in vitro* with peptide-pulsed LPS blasts. CTL activity was assessed by chromium release assay five days after primary *in vitro* stimulation. Retrovirally transduced cells expressing the control antigen



P703P and HLA-A2Kb were used as targets. CTL lines that specifically recognized both p5-pulsed targets as well as P703P-expressing targets were identified.

Human *in vitro* priming experiments demonstrated that the p5 peptide is immunogenic in humans. Dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by culturing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, the DC were pulsed with 1 ug/ml p5 peptide and cultured with CD8+ T cell enriched PBMC. CTL lines were restimulated on a weekly basis with p5-pulsed monocytes. Five to six weeks after initiation of the CTL cultures, CTL recognition of p5-pulsed target cells was demonstrated. CTL were additionally shown to recognize human cells transduced to express P703P, demonstrating that p5 is a naturally processed epitope.

Studies identifying a further peptide epitope (referred to as peptide 4) derived from the prostate tumor-specific antigen P703P that is capable of being recognized by CD4 T cells on the surface of cells in the context of HLA class II molecules were carried out as follows. The amino acid sequence for peptide 4 is provided in SEQ ID NO: 781, with the corresponding cDNA sequence being provided in SEQ ID NO: 782.

Twenty 15-mer peptides overlapping by 10 amino acids and derived from the carboxy-terminal fragment of P703P were generated using standard procedures. Dendritic cells (DC) were derived from PBMC of a normal female donor using GM-CSF and IL-4 by standard protocols. CD4 T cells were generated from the same donor as the DC using MACS beads and negative selection. DC were pulsed overnight with pools of the 15-mer peptides, with each peptide at a final concentration of 0.25 microgram/ml. Pulsed DC were washed and plated at  $1 \times 10^4$  cells/well of 96-well V-bottom plates and purified CD4 T cells were added at  $1 \times 10^5$ /well. Cultures were supplemented with 60 ng/ml IL-6 and 10 ng/ml IL-12 and incubated at 37 °C. Cultures were restimulated as above on a weekly basis using DC generated and pulsed as above as antigen presenting cells, supplemented with 5 ng/ml IL-7 and 10 u/ml IL-2. Following 4 *in vitro* stimulation cycles, 96 lines (each line corresponding to one well) were tested for specific proliferation



and cytokine production in response to the stimulating pools with an irrelevant pool of peptides derived from mammaglobin being used as a control.

One line (referred to as 1-F9) was identified from pool #1 that demonstrated specific proliferation (measured by <sup>3</sup>H proliferation assays) and cytokine production (measured by interferon-gamma ELISA assays) in response to pool #1 of P703P peptides. This line was further tested for specific recognition of the peptide pool, specific recognition of individual peptides in the pool, and in HLA mismatch analyses to identify the relevant restricting allele. Line 1-F9 was found to specifically proliferate and produce interferon-gamma in response to peptide pool #1, and also to peptide 4 (SEQ ID NO: 781). Peptide 4 corresponds to amino acids 126-140 of SEQ ID NO: 327. Peptide titration experiments were conducted to assess the sensitivity of line 1-F9 for the specific peptide. The line was found to specifically respond to peptide 4 at concentrations as low as 0.25 ng/ml, indicating that the T cells are very sensitive and therefore likely to have high affinity for the epitope.

To determine the HLA restriction of the P703P response, a panel of antigen presenting cells (APC) was generated that was partially matched with the donor used to generate the T cells. The APC were pulsed with the peptide and used in proliferation and cytokine assays together with line 1-F9. APC matched with the donor at HLA-DRB0701 and HLA-DQB02 alleles were able to present the peptide to the T cells, indicating that the P703P-specific response is restricted to one of these alleles.

In further studies, twenty-four 15-mer peptides overlapping by 10 amino acids and derived from the N-terminal fragment of P703P (corresponding to amino acids 27-154 of SEQ ID NO: 525) were generated by standard procedures and their ability to be recognized by CD4 cells was determined essentially as described above. DC were pulsed overnight with pools of the peptides with each peptide at a final concentration of 10 microgram/ml. A large number of individual CD4 T cell lines (65/480) demonstrated significant proliferation and cytokine release (IFN-gamma) in response to the P703P peptide pools but not to a control peptide pool. The CD4 T cell lines which demonstrated specific activity were restimulated on the appropriate pool of P703P peptides and reassayed on the individual peptides of each pool as well as a peptide dose titration of the pool of peptides in a IFN-gamma release assay and in a proliferation assay.



Sixteen immunogenic peptides were recognized by the T cells from the entire set of peptide antigens tested. The amino acid sequences of these peptides are provided in SEQ ID NO: 799-814, with the corresponding cDNA sequences being provided in SEQ ID NO: 783-798, respectively. In some cases the peptide reactivity of the T cell line could be mapped to a single peptide, however some could be mapped to more than one peptide in each pool. Those CD4 T cell lines that displayed a representative pattern of recognition from each peptide pool with a reasonable affinity for peptide were chosen for further analysis (I-1A, -6A; II-4C, -5E; III-6E, IV-4B, -3F, -9B, -10F, V-5B, -4D, and -10F). These CD4 T cells lines were restimulated on the appropriate individual peptide and reassayed on autologous DC pulsed with a truncated form of recombinant P703P protein made in *E. coli* (a.a. 96 - 254 of SEQ ID NO: 525), full-length P703P made in the baculovirus expression system, and a fusion between influenza virus NS1 and P703P made in *E. coli*. Of the T cell lines tested, line I-1A recognized specifically the truncated form of P703P (*E. coli*) but no other recombinant form of P703P. This line also recognized the peptide used to elicit the T cells. Line 2-4C recognized the truncated form of P703P (*E. coli*) and the full length form of P703P made in baculovirus, as well as peptide. The remaining T cell lines tested were either peptide-specific only (II-5E, II-6F, IV-4B, IV-3F, IV-9B, IV-10F, V-5B and V-4D) or were non-responsive to any antigen tested (V-10F). These results demonstrate that the peptide sequence RPLLANDLMLIKLDE (SEQ ID NO: 814; corresponding to a.a. 110-124 of SEQ ID NO: 525) recognized by the T cell line I-1A, and the peptide sequences SVSESDTIRSISIAS (SEQ ID NO: 811; corresponding to a.a. 125-139 of SEQ ID NO: 525) and ISIASQCPTAGNSCL (SEQ ID NO: 810; corresponding to a.a. 135-149 of SEQ ID NO: 525) recognized by the T cell line II-4C may be naturally processed epitopes of the P703P protein.

25



EXAMPLE 11  
EXPRESSION OF A BREAST TUMOR-DERIVED ANTIGEN  
IN PROSTATE

5           Isolation of the antigen B305D from breast tumor by differential display is described in US Patent Application No. 08/700,014, filed August 20, 1996. Several different splice forms of this antigen were isolated. The determined cDNA sequences for these splice forms are provided in SEQ ID NO: 366-375, with the predicted amino acid sequences corresponding to the sequences of SEQ ID NO: 292, 298 and 301-303 being  
10 provided in SEQ ID NO: 299-306, respectively. In further studies, a splice variant of the cDNA sequence of SEQ ID NO: 366 was isolated which was found to contain an additional guanine residue at position 884 (SEQ ID NO: 530), leading to a frameshift in the open reading frame. The determined DNA sequence of this ORF is provided in SEQ ID NO: 531. This frameshift generates a protein sequence (provided in SEQ ID NO: 532) of 293  
15 amino acids that contains the C-terminal domain common to the other isoforms of B305D but that differs in the N-terminal region.

          The expression levels of B305D in a variety of tumor and normal tissues were examined by real time PCR and by Northern analysis. The results indicated that B305D is highly expressed in breast tumor, prostate tumor, normal prostate and normal  
20 testes, with expression being low or undetectable in all other tissues examined (colon tumor, lung tumor, ovary tumor, and normal bone marrow, colon, kidney, liver, lung, ovary, skin, small intestine, stomach). Using real-time PCR on a panel of prostate tumors, expression of B305D in prostate tumors was shown to increase with increasing Gleason grade, demonstrating that expression of B305D increases as prostate cancer progresses.

25



## EXAMPLE 12

GENERATION OF HUMAN CTL *IN VITRO* USING WHOLE GENE PRIMING AND STIMULATION  
TECHNIQUES WITH PROSTATE-SPECIFIC ANTIGEN

- 5                   Using *in vitro* whole-gene priming with P501S-vaccinia infected DC (see, for example, Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines were derived that specifically recognize autologous fibroblasts transduced with P501S (also known as L1-12), as determined by interferon- $\gamma$  ELISPOT analysis as described above. Using a panel of HLA-mismatched B-LCL lines transduced with P501S, these CTL
- 10 lines were shown to be likely restricted to HLAB class I allele. Specifically, dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC were infected overnight with recombinant P501S vaccinia virus at a multiplicity of infection (M.O.I) of
- 15 five, and matured overnight by the addition of 3  $\mu$ g/ml CD40 ligand. Virus was inactivated by UV irradiation. CD8<sup>+</sup> T cells were isolated using a magnetic bead system, and priming cultures were initiated using standard culture techniques. Cultures were restimulated every 7-10 days using autologous primary fibroblasts retrovirally transduced with P501S and CD80. Following four stimulation cycles, CD8<sup>+</sup> T cell lines were identified that
- 20 specifically produced interferon- $\gamma$  when stimulated with P501S and CD80-transduced autologous fibroblasts. A panel of HLA-mismatched B-LCL lines transduced with P501S were generated to define the restriction allele of the response. By measuring interferon- $\gamma$  in an ELISPOT assay, the P501S specific response was shown to be likely restricted by HLA B alleles. These results demonstrate that a CD8<sup>+</sup> CTL response to P501S can be elicited.
- 25                   To identify the epitope(s) recognized, cDNA encoding P501S was fragmented by various restriction digests, and sub-cloned into the retroviral expression vector pBIB-KS. Retroviral supernatants were generated by transfection of the helper packaging line Phoenix-Ampho. Supernatants were then used to transduce Jurkat/A2Kb cells for CTL screening. CTL were screened in IFN-gamma ELISPOT assays against these



A2Kb targets transduced with the “library” of P501S fragments. Initial positive fragments P501S/H3 and P501S/F2 were sequenced and found to encode amino acids 106-553 and amino acids 136-547, respectively, of SEQ ID NO: 113. A truncation of H3 was made to encode amino acid residues 106-351 of SEQ ID NO: 113, which was unable to stimulate the CTL, thus localizing the epitope to amino acid residues 351-547. Additional fragments encoding amino acids 1-472 (Fragment A) and amino acids 1-351 (Fragment B) were also constructed. Fragment A but not Fragment B stimulated the CTL thus localizing the epitope to amino acid residues 351-472. Overlapping 20-mer and 18-mer peptides representing this region were tested by pulsing Jurkat/A2Kb cells versus CTL in an IFN-  
 5 gamma assay. Only peptides P501S-369(20) and P501S-369(18) stimulated the CTL. Nine-mer and 10-mer peptides representing this region were synthesized and similarly tested. Peptide P501S-370 (SEQ ID NO: 539) was the minimal 9-mer giving a strong response. Peptide P501S-376 (SEQ ID NO: 540) also gave a weak response, suggesting that it might represent a cross-reactive epitope.

15 In subsequent studies, the ability of primary human B cells transduced with P501S to prime MHC class I-restricted, P501S-specific, autologous CD8 T cells was examined. Primary B cells were derived from PBMC of a homozygous HLA-A2 donor by culture in CD40 ligand and IL-4, transduced at high frequency with recombinant P501S in the vector pBIB, and selected with blastocidin-S. For *in vitro* priming, purified CD8+ T  
 20 cells were cultured with autologous CD40 ligand + IL-4 derived, P501S-transduced B cells in a 96-well microculture format. These CTL microcultures were re-stimulated with P501S-transduced B cells and then assayed for specificity. Following this initial screen, microcultures with significant signal above background were cloned on autologous EBV-transformed B cells (BLCL), also transduced with P501S. Using IFN-gamma ELISPOT  
 25 for detection, several of these CD8 T cell clones were found to be specific for P501S, as demonstrated by reactivity to BLCL/P501S but not BLCL transduced with control antigen. It was further demonstrated that the anti-P501S CD8 T cell specificity is HLA-A2-restricted. First, antibody blocking experiments with anti-HLA-A,B,C monoclonal antibody (W6.32), anti-HLA-B,C monoclonal antibody (B1.23.2) and a control monoclonal



antibody showed that only the anti-HLA-A,B,C antibody blocked recognition of P501S-expressing autologous BLCL. Secondly, the anti-P501S CTL also recognized an HLA-A2 matched, heterologous BLCL transduced with P501S, but not the corresponding EGFP transduced control BLCL.

5

### EXAMPLE 13

#### IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS

##### BY MICROARRAY ANALYSIS

10                This Example describes the isolation of certain prostate-specific polypeptides from a prostate tumor cDNA library.

                 A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate  
 15    normal tissues (not including testis). 372 clones were identified, and 319 were successfully sequenced. Table I presents a summary of these clones, which are shown in SEQ ID NOs:385-400. Of these sequences SEQ ID NOs:386, 389, 390 and 392 correspond to novel genes, and SEQ ID NOs: 393 and 396 correspond to previously identified sequences. The others (SEQ ID NOs:385, 387, 388, 391, 394, 395 and 397-400) correspond to known  
 20    sequences, as shown in Table I.

#### Table I



### Summary of Prostate Tumor Antigens

Known Genes	Previously Identified Genes	Novel Genes
T-cell gamma chain	P504S	23379 (SEQ ID NO:389)
Kallikrein	P1000C	23399 (SEQ ID NO:392)
Vector	P501S	23320 (SEQ ID NO:386)
CGI-82 protein mRNA (23319; SEQ ID NO:385)	P503S	23381 (SEQ ID NO:390)
PSA	P510S	
Ald. 6 Dehyd.	P784P	
L-itol-2 dehydrogenase (23376; SEQ ID NO:388)	P502S	
Ets transcription factor PDEF (22672; SEQ ID NO:398)	P706P	
hTGR (22678; SEQ ID NO:399)	19142.2, bangur.seq (22621; SEQ ID NO:396)	
KIAA0295(22685; SEQ ID NO:400)	5566.1 Wang (23404; SEQ ID NO:393)	
Prostatic Acid Phosphatase(22655; SEQ ID NO:397)	P712P	
transglutaminase (22611; SEQ ID NO:395)	P778P	
HDLBP (23508; SEQ ID NO:394)		
CGI-69 Protein(23367; SEQ ID NO:387)		
KIAA0122(23383; SEQ ID NO:391)		
TEEG		

CGI-82 showed 4.06 fold over-expression in prostate tissues as compared to  
5 other normal tissues tested. It was over-expressed in 43% of prostate tumors, 25% normal



prostate, not detected in other normal tissues tested. L-idoitol-2 dehydrogenase showed 4.94 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 90% of prostate tumors, 100% of normal prostate, and not detected in other normal tissues tested. Ets transcription factor PDEF showed 5.55 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% prostate tumors, 25% normal prostate and not detected in other normal tissues tested. hTGR1 showed 9.11 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 63% of prostate tumors and is not detected in normal tissues tested including normal prostate. KIAA0295 showed 5.59 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% of prostate tumors, low to undetectable in normal tissues tested including normal prostate tissues. Prostatic acid phosphatase showed 9.14 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 67% of prostate tumors, 50% of normal prostate, and not detected in other normal tissues tested. Transglutaminase showed 14.84 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 30% of prostate tumors, 50% of normal prostate, and is not detected in other normal tissues tested. High density lipoprotein binding protein (HDLBP) showed 28.06 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% of normal prostate, and is undetectable in all other normal tissues tested. CGI-69 showed 3.56 fold over-expression in prostate tissues as compared to other normal tissues tested. It is a low abundant gene, detected in more than 90% of prostate tumors, and in 75% normal prostate tissues. The expression of this gene in normal tissues was very low. KIAA0122 showed 4.24 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 57% of prostate tumors, it was undetectable in all normal tissues tested including normal prostate tissues. 19142.2 bangur showed 23.25 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors and 100% of normal prostate. It was undetectable in other normal tissues tested. 5566.1 Wang showed 3.31 fold over-



expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% normal prostate and was also over-expressed in normal bone marrow, pancreas, and activated PBMC. Novel clone 23379 (also referred to as P553S) showed 4.86 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in 97% of prostate tumors and 75% normal prostate and is undetectable in all other normal tissues tested. Novel clone 23399 showed 4.09 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 27% of prostate tumors and was undetectable in all normal tissues tested including normal prostate tissues. Novel clone 23320 showed 3.15 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in all prostate tumors and 50% of normal prostate tissues. It was also expressed in normal colon and trachea. Other normal tissues do not express this gene at high level.

Subsequent full-length cloning studies on P553S, using standard techniques, revealed that this clone is an incomplete spliced form of P501S. The determined cDNA sequences for four splice variants of P553S are provided in SEQ ID NO: 702-705. An amino acid sequence encoded by SEQ ID NO: 705 is provided in SEQ ID NO: 706. The cDNA sequence of SEQ ID NO: 702 was found to contain two open reading frames (ORFs). The amino acid sequences encoded by these two ORFs are provided in SEQ ID NO: 707 and 708.

#### EXAMPLE 14

##### IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY ELECTRONIC SUBTRACTION

This Example describes the use of an electronic subtraction technique to identify prostate-specific antigens.

Potential prostate-specific genes present in the GenBank human EST database were identified by electronic subtraction (similar to that described by Vasmatizis et al., *Proc. Natl. Acad. Sci. USA* 95:300-304, 1998). The sequences of EST clones



(43,482) derived from various prostate libraries were obtained from the GenBank public human EST database. Each prostate EST sequence was used as a query sequence in a BLASTN (National Center for Biotechnology Information) search against the human EST database. All matches considered identical (length of matching sequence >100 base pairs, density of identical matches over this region > 70%) were grouped (aligned) together in a cluster. Clusters containing more than 200 ESTs were discarded since they probably represented repetitive elements or highly expressed genes such as those for ribosomal proteins. If two or more clusters shared common ESTs, those clusters were grouped together into a “supercluster,” resulting in 4,345 prostate superclusters.

Records for the 479 human cDNA libraries represented in the GenBank release were downloaded to create a database of these cDNA library records. These 479 cDNA libraries were grouped into three groups: Plus (normal prostate and prostate tumor libraries, and breast cell line libraries, in which expression was desired), Minus (libraries from other normal adult tissues, in which expression was not desirable), and Other (libraries from fetal tissue, infant tissue, tissues found only in women, non-prostate tumors and cell lines other than prostate cell lines, in which expression was considered to be irrelevant). A summary of these library groups is presented in Table II.

Table II

Prostate cDNA Libraries and ESTs

Library	# of Libraries	# of ESTs
Plus	25	43,482
Normal	11	18,875
Tumor	11	21,769
Cell lines	3	2,838
Minus	166	
Other	287	



Each supercluster was analyzed in terms of the ESTs within the supercluster. The tissue source of each EST clone was noted and used to classify the superclusters into four groups: Type 1- EST clones found in the Plus group libraries only; no expression detected in Minus or Other group libraries; Type 2- EST clones derived from the Plus and Other group libraries only; no expression detected in the Minus group; Type 3- EST clones derived from the Plus, Minus and Other group libraries, but the number of ESTs derived from the Plus group is higher than in either the Minus or Other groups; and Type 4- EST clones derived from Plus, Minus and Other group libraries, but the number derived from the Plus group is higher than the number derived from the Minus group. This analysis identified 4,345 breast clusters (*see* Table III). From these clusters, 3,172 EST clones were ordered from Research Genetics, Inc., and were received as frozen glycerol stocks in 96-well plates.

Table III

Prostate Cluster Summary

Type	# of Superclusters	# of ESTs Ordered
1	688	677
2	2899	2484
3	85	11
4	673	0
Total	4345	3172

The EST clone inserts were PCR-amplified using amino-linked PCR primers for Synteni microarray analysis. When more than one PCR product was obtained for a particular clone, that PCR product was not used for expression analysis. In total, 2,528 clones from the electronic subtraction method were analyzed by microarray analysis to identify electronic subtraction breast clones that had high levels of tumor vs. normal



tissue mRNA. Such screens were performed using a Synteni (Palo Alto, CA) microarray, according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Within these analyses, the clones were arrayed on the chip, which was then probed with fluorescent probes generated from normal and tumor prostate cDNA, as well as various other normal tissues. The slides were scanned and the fluorescence intensity was measured.

Clones with an expression ratio greater than 3 (*i.e.*, the level in prostate tumor and normal prostate mRNA was at least three times the level in other normal tissue mRNA) were identified as prostate tumor-specific sequences (Table IV). The sequences of these clones are provided in SEQ ID NO: 401-453, with certain novel sequences shown in SEQ ID NO: 407, 413, 416-419, 422, 426, 427 and 450.

Table IV

Prostate-tumor Specific Clones

SEQ ID NO.	Sequence Designation	Comments
401	22545	previously identified P1000C
402	22547	previously identified P704P
403	22548	known
404	22550	known
405	22551	PSA
406	22552	prostate secretory protein 94
407	22553	novel
408	22558	previously identified P509S
409	22562	glandular kallikrein
410	22565	previously identified P1000C
411	22567	PAP
412	22568	B1006C (breast tumor antigen)
413	22570	novel
414	22571	PSA
415	22572	previously identified P706P
416	22573	novel
417	22574	novel



418	22575	novel
419	22580	novel
420	22581	PAP
421	22582	prostatic secretory protein 94
422	22583	novel
423	22584	prostatic secretory protein 94
424	22585	prostatic secretory protein 94
425	22586	known
426	22587	novel
427	22588	novel
428	22589	PAP
429	22590	known
430	22591	PSA
431	22592	known
432	22593	Previously identified P777P
433	22594	T cell receptor gamma chain
434	22595	Previously identified P705P
435	22596	Previously identified P707P
436	22847	PAP
437	22848	known
438	22849	prostatic secretory protein 57
439	22851	PAP
440	22852	PAP
441	22853	PAP
442	22854	previously identified P509S
443	22855	previously identified P705P
444	22856	previously identified P774P
445	22857	PSA
446	23601	previously identified P777P
447	23602	PSA
448	23605	PSA
449	23606	PSA
450	23612	novel
451	23614	PSA
452	23618	previously identified P1000C
453	23622	previously identified P705P

Further studies on the clone of SEQ ID NO: 407 (also referred to as P1020C) led to the isolation of an extended cDNA sequence provided in SEQ ID NO: 591. This extended cDNA sequence was found to contain an open reading frame that encodes



the predicted amino acid sequence of SEQ ID NO: 592. The P1020C cDNA and amino acid sequences were found to show some similarity to the human endogenous retroviral HERV-K pol gene and protein.

5

## EXAMPLE 15

## FURTHER IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY MICROARRAY ANALYSIS

This Example describes the isolation of additional prostate-specific polypeptides from a prostate tumor cDNA library.

10

A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 142 clones were identified and sequenced. Certain of these clones are shown in SEQ ID NO: 454-467. Of these sequences, SEQ ID NO: 459-15 461 represent novel genes. The others (SEQ ID NO: 454-458 and 461-467) correspond to known sequences.

## EXAMPLE 16

## FURTHER CHARACTERIZATION OF PROSTATE-SPECIFIC ANTIGEN P710P

20

This Example describes the full length cloning of P710P.

The prostate cDNA library described above was screened with the P710P fragment described above. One million colonies were plated on LB/Ampicillin plates. Nylon membrane filters were used to lift these colonies, and the cDNAs picked up by these 25 filters were then denatured and cross-linked to the filters by UV light. The P710P fragment was radiolabeled and used to hybridize with the filters. Positive cDNA clones were selected and their cDNAs recovered and sequenced by an automatic Perkin Elmer/Applied Biosystems Division Sequencer. Four sequences were obtained, and are presented in SEQ ID NO: 468-471. These sequences appear to represent different splice variants of the



P710P gene. Subsequent comparison of the cDNA sequences of P710P with those in Genbank revealed homology to the DD3 gene (Genbank accession numbers AF103907 & AF103908). The cDNA sequence of DD3 is provided in SEQ ID NO: 690.

5

## EXAMPLE 17

## PROTEIN EXPRESSION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

This example describes the expression and purification of the prostate-specific antigen P501S in *E. coli*, baculovirus and mammalian cells.

10 **A) EXPRESSION IN *E. COLI***

Expression of the full-length form of P501S was attempted by first cloning P501S without the leader sequence (amino acids 36-553 of SEQ ID NO: 113) downstream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) in pET17b. Specifically, P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW003 (SEQ ID NO: 486). AW025 is a sense cloning primer that  
15 contains a HindIII site. AW003 is an antisense cloning primer that contains an EcoRI site. DNA amplification was performed using 5 µl 10X Pfu buffer, 1 µl 20 mM dNTPs, 1 µl each of the PCR primers at 10 µM concentration, 40 µl water, 1 µl Pfu DNA polymerase (Stratagene, La Jolla, CA) and 1 µl DNA at 100 ng/µl. Denaturation at 95°C was  
20 performed for 30 sec, followed by 10 cycles of 95°C for 30 sec, 60°C for 1 min and by 72°C for 3 min. 20 cycles of 95°C for 30 sec, 65°C for 1 min and by 72°C for 3 min, and lastly by 1 cycle of 72°C for 10 min. The PCR product was cloned to Ra12m/pET17b using HindIII and EcoRI. The sequence of the resulting fusion construct (referred to as Ra12-P501S-F) was confirmed by DNA sequencing.

25 The fusion construct was transformed into BL21(DE3)pLysE, pLysS and CodonPlus *E. coli* (Stratagene) and grown overnight in LB broth with kanamycin. The resulting culture was induced with IPTG. Protein was transferred to PVDF membrane and blocked with 5% non-fat milk (in PBS-Tween buffer), washed three times and incubated



with mouse anti-His tag antibody (Clontech) for 1 hour. The membrane was washed 3 times and probed with HRP-Protein A (Zymed) for 30 min. Finally, the membrane was washed 3 times and developed with ECL (Amersham). No expression was detected by Western blot. Similarly, no expression was detected by Western blot when the Ra12-  
 5 P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage (Invitrogen).

An N-terminal fragment of P501S (amino acids 36-325 of SEQ ID NO: 113) was cloned down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 in pET17b as follows. P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW027 (SEQ ID NO: 487). AW027 is an antisense cloning primer that  
 10 contains an EcoRI site and a stop codon. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The fusion construct (referred to as Ra12-P501S-N) was confirmed by DNA sequencing.

The Ra12-P501S-N fusion construct was used for expression in  
 15 BL21(DE3)pLysE, pLysS and CodonPlus, essentially as described above. Using Western blot analysis, protein bands were observed at the expected molecular weight of 36 kDa. Some high molecular weight bands were also observed, probably due to aggregation of the recombinant protein. No expression was detected by Western blot when the Ra12-P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage.

A fusion construct comprising a C-terminal portion of P501S (amino acids 257-553 of SEQ ID NO: 113) located down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) was prepared as follows. P501S DNA was used to perform PCR using the primers AW026 (SEQ ID NO: 488) and AW003 (SEQ ID NO: 486). AW026 is a sense cloning primer that contains a HindIII site. DNA  
 25 amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The sequence for the fusion construct (referred to as Ra12-P501S-C) was confirmed.

The Ra12-P501S-C fusion construct was used for expression in BL21(DE3)pLysE, pLysS and CodonPlus, as described above. A small amount of protein



was detected by Western blot, with some molecular weight aggregates also being observed. Expression was also detected by Western blot when the Ra12-P501S-C fusion was used for expression in BL21CodonPlus induced by CE6 phage.

#### **B) EXPRESSION OF P501S IN BACULOVIRUS**

5                   The Bac-to-Bac baculovirus expression system (BRL Life Technologies, Inc.) was used to express P501S protein in insect cells. Full-length P501S (SEQ ID NO: 113) was amplified by PCR and cloned into the XbaI site of the donor plasmid pFastBacI. The recombinant bacmid and baculovirus were prepared according to the manufacturer's instructions. The recombinant baculovirus was amplified in Sf9 cells and the high titer  
10   viral stocks were utilized to infect High Five cells (Invitrogen) to make the recombinant protein. The identity of the full-length protein was confirmed by N-terminal sequencing of the recombinant protein and by Western blot analysis (Figure 7). Specifically, 0.6 million High Five cells in 6-well plates were infected with either the unrelated control virus BV/ECD\_PD (lane 2), with recombinant baculovirus for P501S at different amounts or  
15   MOIs (lanes 4-8), or were uninfected (lane 3). Cell lysates were run on SDS-PAGE under reducing conditions and analyzed by Western blot with the anti-P501S monoclonal antibody P501S-10E3-G4D3 (prepared as described below). Lane 1 is the biotinylated protein molecular weight marker (BioLabs).

                  The localization of recombinant P501S in the insect cells was investigated  
20   as follows. The insect cells overexpressing P501S were fractionated into fractions of nucleus, mitochondria, membrane and cytosol. Equal amounts of protein from each fraction were analyzed by Western blot with a monoclonal antibody against P501S. Due to the scheme of fractionation, both nucleus and mitochondria fractions contain some plasma membrane components. However, the membrane fraction is basically free from  
25   mitochondria and nucleus. P501S was found to be present in all fractions that contain the membrane component, suggesting that P501S may be associated with plasma membrane of the insect cells expressing the recombinant protein.



### **c) EXPRESSION OF P501S IN MAMMALIAN CELLS**

Full-length P501S (553AA) was cloned into various mammalian expression vectors, including pCEP4 (Invitrogen), pVR1012 (Vical, San Diego, CA) and a modified form of the retroviral vector pBMN, referred to as pBIB. Transfection of P501S/pCEP4 and P501S/pVR1012 into HEK293 fibroblasts was carried out using the Fugene transfection reagent (Boehringer Mannheim). Briefly, 2  $\mu$ l of Fugene reagent was diluted into 100  $\mu$ l of serum-free media and incubated at room temperature for 5-10 min. This mixture was added to 1  $\mu$ g of P501S plasmid DNA, mixed briefly and incubated for 30 minutes at room temperature. The Fugene/DNA mixture was added to cells and incubated for 24-48 hours. Expression of recombinant P501S in transfected HEK293 fibroblasts was detected by means of Western blot employing a monoclonal antibody to P501S.

Transfection of p501S/pCEP4 into CHO-K cells (American Type Culture Collection, Rockville, MD) was carried out using GenePorter transfection reagent (Gene Therapy Systems, San Diego, CA). Briefly, 15  $\mu$ l of GenePorter was diluted in 500  $\mu$ l of serum-free media and incubated at room temperature for 10 min. The GenePorter/media mixture was added to 2  $\mu$ g of plasmid DNA that was diluted in 500  $\mu$ l of serum-free media, mixed briefly and incubated for 30 min at room temperature. CHO-K cells were rinsed in PBS to remove serum proteins, and the GenePorter/DNA mix was added and incubated for 5 hours. The transfected cells were then fed an equal volume of 2x media and incubated for 24-48 hours.

FACS analysis of P501S transiently infected CHO-K cells, demonstrated surface expression of P501S. Expression was detected using rabbit polyclonal antisera raised against a P501S peptide, as described below. Flow cytometric analysis was performed using a FaCScan (Becton Dickinson), and the data were analyzed using the Cell Quest program.



## EXAMPLE 18

PREPARATION AND CHARACTERIZATION OF ANTIBODIES  
AGAINST PROSTATE-SPECIFIC POLYPEPTIDES

5    **A) PREPARATION AND CHARACTERIZATION OF POLYCLONAL ANTIBODIES AGAINST  
P703P, P504S AND P509S**

Polyclonal antibodies against P703P, P504S and P509S were prepared as follows.

Each prostate tumor antigen expressed in an *E. coli* recombinant expression  
10    system was grown overnight in LB broth with the appropriate antibiotics at 37°C in a  
shaking incubator. The next morning, 10 ml of the overnight culture was added to 500 ml  
to 2x YT plus appropriate antibiotics in a 2L-baffled Erlenmeyer flask. When the Optical  
Density (at 560 nm) of the culture reached 0.4-0.6, the cells were induced with IPTG (1  
mM). Four hours after induction with IPTG, the cells were harvested by centrifugation.  
15    The cells were then washed with phosphate buffered saline and centrifuged again. The  
supernatant was discarded and the cells were either frozen for future use or immediately  
processed. Twenty ml of lysis buffer was added to the cell pellets and vortexed. To break  
open the *E. coli* cells, this mixture was then run through the French Press at a pressure of  
16,000 psi. The cells were then centrifuged again and the supernatant and pellet were  
20    checked by SDS-PAGE for the partitioning of the recombinant protein. For proteins that  
localized to the cell pellet, the pellet was resuspended in 10 mM Tris pH 8.0, 1% CHAPS  
and the inclusion body pellet was washed and centrifuged again. This procedure was  
repeated twice more. The washed inclusion body pellet was solubilized with either 8 M  
urea or 6 M guanidine HCl containing 10 mM Tris pH 8.0 plus 10 mM imidazole. The  
25    solubilized protein was added to 5 ml of nickel-chelate resin (Qiagen) and incubated for 45  
min to 1 hour at room temperature with continuous agitation. After incubation, the resin  
and protein mixture were poured through a disposable column and the flow through was  
collected. The column was then washed with 10-20 column volumes of the solubilization



buffer. The antigen was then eluted from the column using 8M urea, 10 mM Tris pH 8.0 and 300 mM imidazole and collected in 3 ml fractions. A SDS-PAGE gel was run to determine which fractions to pool for further purification.

As a final purification step, a strong anion exchange resin such as HiPrepQ (Biorad) was equilibrated with the appropriate buffer and the pooled fractions from above were loaded onto the column. Each antigen was eluted off the column with a increasing salt gradient. Fractions were collected as the column was run and another SDS-PAGE gel was run to determine which fractions from the column to pool. The pooled fractions were dialyzed against 10 mM Tris pH 8.0. The proteins were then vialled after filtration through a 0.22 micron filter and the antigens were frozen until needed for immunization.

Four hundred micrograms of each prostate antigen was combined with 100 micrograms of muramyl dipeptide (MDP). Every four weeks rabbits were boosted with 100 micrograms mixed with an equal volume of Incomplete Freund's Adjuvant (IFA). Seven days following each boost, the animal was bled. Sera was generated by incubating the blood at 4°C for 12-4 hours followed by centrifugation.

Ninety-six well plates were coated with antigen by incubating with 50 microliters (typically 1 microgram) of recombinant protein at 4 °C for 20 hours. 250 microliters of BSA blocking buffer was added to the wells and incubated at room temperature for 2 hours. Plates were washed 6 times with PBS/0.01% Tween. Rabbit sera was diluted in PBS. Fifty microliters of diluted sera was added to each well and incubated at room temperature for 30 min. Plates were washed as described above before 50 microliters of goat anti-rabbit horse radish peroxidase (HRP) at a 1:10000 dilution was added and incubated at room temperature for 30 min. Plates were again washed as described above and 100 microliters of TMB microwell peroxidase substrate was added to each well. Following a 15 min incubation in the dark at room temperature, the colorimetric reaction was stopped with 100 microliters of 1N H<sub>2</sub>SO<sub>4</sub> and read immediately at 450 nm. All polyclonal antibodies showed immunoreactivity to the appropriate antigen.



## B) PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST P501S

A murine monoclonal antibody directed against the carboxy-terminus of the prostate-specific antigen P501S was prepared as follows.

A truncated fragment of P501S (amino acids 355-526 of SEQ ID NO: 113) was generated and cloned into the pET28b vector (Novagen) and expressed in *E. coli* as a thioredoxin fusion protein with a histidine tag. The trx-P501S fusion protein was purified by nickel chromatography, digested with thrombin to remove the trx fragment and further purified by an acid precipitation procedure followed by reverse phase HPLC.

Mice were immunized with truncated P501S protein. Serum bleeds from mice that potentially contained anti-P501S polyclonal sera were tested for P501S-specific reactivity using ELISA assays with purified P501S and trx-P501S proteins. Serum bleeds that appeared to react specifically with P501S were then screened for P501S reactivity by Western analysis. Mice that contained a P501S-specific antibody component were sacrificed and spleen cells were used to generate anti-P501S antibody producing hybridomas using standard techniques. Hybridoma supernatants were tested for P501S-specific reactivity initially by ELISA, and subsequently by FACS analysis of reactivity with P501S transduced cells. Based on these results, a monoclonal hybridoma referred to as 10E3 was chosen for further subcloning. A number of subclones were generated, tested for specific reactivity to P501S using ELISA and typed for IgG isotype. The results of this analysis are shown below in Table V. Of the 16 subclones tested, the monoclonal antibody 10E3-G4-D3 was selected for further study.

Table V

### Isotype analysis of murine anti-P501S monoclonal antibodies

Hybridoma clone	Isotype	Estimated [Ig] in supernatant (µg/ml)
4D11	IgG1	14.6
1G1	IgG1	0.6
4F6	IgG1	72
4H5	IgG1	13.8
4H5-E12	IgG1	10.7



Hybridoma clone	Isotype	Estimated [Ig] in supernatant (µg/ml)
4H5-EH2	IgG1	9.2
4H5-H2-A10	IgG1	10
4H5-H2-A3	IgG1	12.8
4H5-H2-A10-G6	IgG1	13.6
4H5-H2-B11	IgG1	12.3
10E3	IgG2a	3.4
10E3-D4	IgG2a	3.8
10E3-D4-G3	IgG2a	9.5
10E3-D4-G6	IgG2a	10.4
10E3-E7	IgG2a	6.5
8H12	IgG2a	0.6

The specificity of 10E3-G4-D3 for P501S was examined by FACS analysis. Specifically, cells were fixed (2% formaldehyde, 10 minutes), permeabilized (0.1% saponin, 10 minutes) and stained with 10E3-G4-D3 at 0.5 – 1 µg/ml, followed by  
5 incubation with a secondary, FITC-conjugated goat anti-mouse Ig antibody (Pharmingen, San Diego, CA). Cells were then analyzed for FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. For analysis of infected cells, B-LCL were infected with a vaccinia vector that expresses P501S. To demonstrate specificity in these assays, B-  
10 LCL transduced with a different antigen (P703P) and uninfected B-LCL vectors were utilized. 10E3-G4-D3 was shown to bind with P501S-transduced B-LCL and also with P501S-infected B-LCL, but not with either uninfected cells or P703P-transduced cells.

To determine whether the epitope recognized by 10E3-G4-D3 was found on the surface or in an intracellular compartment of cells, B-LCL were transduced with P501S  
15 or HLA-B8 as a control antigen and either fixed and permeabilized as described above or directly stained with 10E3-G4-D3 and analyzed as above. Specific recognition of P501S by 10E3-G4-D3 was found to require permeabilization, suggesting that the epitope recognized by this antibody is intracellular.

The reactivity of 10E3-G4-D3 with the three prostate tumor cell lines  
20 Lncap, PC-3 and DU-145, which are known to express high, medium and very low levels of P501S, respectively, was examined by permeabilizing the cells and treating them as



described above. Higher reactivity of 10E3-G4-D3 was seen with Lncap than with PC-3, which in turn showed higher reactivity than DU-145. These results are in agreement with the real time PCR and demonstrate that the antibody specifically recognizes P501S in these tumor cell lines and that the epitope recognized in prostate tumor cell lines is also  
 5 intracellular.

Specificity of 10E3-G4-D3 for P501S was also demonstrated by Western blot analysis. Lysates from the prostate tumor cell lines Lncap, DU-145 and PC-3, from P501S-transiently transfected HEK293 cells, and from non-transfected HEK293 cells were generated. Western blot analysis of these lysates with 10E3-G4-D3 revealed a 46 kDa  
 10 immunoreactive band in Lncap, PC-3 and P501S-transfected HEK cells, but not in DU-145 cells or non-transfected HEK293 cells. P501S mRNA expression is consistent with these results since semi-quantitative PCR analysis revealed that P501S mRNA is expressed in Lncap, to a lesser but detectable level in PC-3 and not at all in DU-145 cells. Bacterially expressed and purified recombinant P501S (referred to as P501SStr2) was recognized by  
 15 10E3-G4-D3 (24 kDa), as was full-length P501S that was transiently expressed in HEK293 cells using either the expression vector VR1012 or pCEP4. Although the predicted molecular weight of P501S is 60.5 kDa, both transfected and “native” P501S run at a slightly lower mobility due to its hydrophobic nature.

Immunohistochemical analysis was performed on prostate tumor and a  
 20 panel of normal tissue sections (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). Tissue samples were fixed in formalin solution for 24 hours and embedded in paraffin before being sliced into 10 micron sections. Tissue sections were permeabilized and incubated with 10E3-G4-D3 antibody for 1 hr. HRP-labeled anti-mouse followed by  
 25 incubation with DAB chromogen was used to visualize P501S immunoreactivity. P501S was found to be highly expressed in both normal prostate and prostate tumor tissue but was not detected in any of the other tissues tested.

To identify the epitope recognized by 10E3-G4-D3, an epitope mapping approach was pursued. A series of 13 overlapping 20-21 mers (5 amino acid overlap; SEQ



ID NO: 489-501) was synthesized that spanned the fragment of P501S used to generate 10E3-G4-D3. Flat bottom 96 well microtiter plates were coated with either the peptides or the P501S fragment used to immunize mice, at 1 microgram/ml for 2 hours at 37 °C. Wells were then aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature, and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified antibody 10E3-G4-D3 was added at 2 fold dilutions (1000 ng – 16 ng) in PBST and incubated for 30 minutes at room temperature. This was followed by washing 6 times with PBST and subsequently incubating with HRP-conjugated donkey anti-mouse IgG (H+L)Affinipure F(ab') fragment (Jackson ImmunoResearch, West Grove, PA) at 1:20000 for 30 minutes. Plates were then washed and incubated for 15 minutes in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 8, reactivity was seen with the peptide of SEQ ID NO: 496 (corresponding to amino acids 439-459 of P501S) and with the P501S fragment but not with the remaining peptides, demonstrating that the epitope recognized by 10E3-G4-D3 is localized to amino acids 439-459 of SEQ ID NO: 113.

In order to further evaluate the tissue specificity of P501S, multi-array immunohistochemical analysis was performed on approximately 4700 different human tissues encompassing all the major normal organs as well as neoplasias derived from these tissues. Sixty-five of these human tissue samples were of prostate origin. Tissue sections 0.6 mm in diameter were formalin-fixed and paraffin embedded. Samples were pretreated with HIER using 10 mM citrate buffer pH 6.0 and boiling for 10 min. Sections were stained with 10E3-G4-D3 and P501S immunoreactivity was visualized with HRP. All the 65 prostate tissues samples (5 normal, 55 untreated prostate tumors, 5 hormone refractory prostate tumors) were positive, showing distinct perinuclear staining. All other tissues examined were negative for P501S expression.



**c) PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST P503S**

A fragment of P503S (amino acids 113-241 of SEQ ID NO: 114) was expressed and purified from bacteria essentially as described above for P501S and used to immunize both rabbits and mice. Mouse monoclonal antibodies were isolated using  
 5 standard hybridoma technology as described above. Rabbit monoclonal antibodies were isolated using Selected Lymphocyte Antibody Method (SLAM) technology at Immgenics Pharmaceuticals (Vancouver, BC, Canada). Table VI, below, lists the monoclonal antibodies that were developed against P503S.

10

Table VI

Antibody	Species
20D4	Rabbit
JA1	Rabbit
1A4	Mouse
1C3	Mouse
1C9	Mouse
1D12	Mouse
2A11	Mouse
2H9	Mouse
4H7	Mouse
8A8	Mouse
8D10	Mouse
9C12	Mouse
6D12	Mouse

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 20D4 and JA1 were determined and are  
 15 provided in SEQ ID NO: 502 and 503, respectively.

In order to better define the epitope binding region of each of the antibodies, a series of overlapping peptides were generated that span amino acids 109-213 of SEQ ID NO: 114. These peptides were used to epitope map the anti-P503S monoclonal antibodies by ELISA as follows. The recombinant fragment of P503S that was employed as the



immunogen was used as a positive control. Ninety-six well microtiter plates were coated with either peptide or recombinant antigen at 20 ng/well overnight at 4 °C. Plates were aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature then washed in PBS containing 0.1% Tween 20 (PBST). Purified  
 5 rabbit monoclonal antibodies diluted in PBST were added to the wells and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and incubation with Protein-A HRP conjugate at a 1:2000 dilution for a further 30 min. Plates were washed six times in PBST and incubated with tetramethylbenzidine (TMB) substrate for a further 15 min. The reaction was stopped by the addition of 1N sulfuric acid and  
 10 plates were read at 450 nm using at ELISA plate reader. ELISA with the mouse monoclonal antibodies was performed with supernatants from tissue culture run neat in the assay.

All of the antibodies bound to the recombinant P503S fragment, with the exception of the negative control SP2 supernatant. 20D4, JA1 and 1D12 bound strictly to  
 15 peptide #2101 (SEQ ID NO: 504), which corresponds to amino acids 151-169 of SEQ ID NO: 114. 1C3 bound to peptide #2102 (SEQ ID NO: 505), which corresponds to amino acids 165-184 of SEQ ID NO: 114. 9C12 bound to peptide #2099 (SEQ ID NO: 522), which corresponds to amino acids 120-139 of SEQ ID NO: 114. The other antibodies bind to regions that were not examined in these studies.

20 Subsequent to epitope mapping, the antibodies were tested by FACS analysis on a cell line that stably expressed P503S to confirm that the antibodies bind to cell surface epitopes. Cells stably transfected with a control plasmid were employed as a negative control. Cells were stained live with no fixative. 0.5 ug of anti-P503S monoclonal antibody was added and cells were incubated on ice for 30 min before being  
 25 washed twice and incubated with a FITC-labelled goat anti-rabbit or mouse secondary antibody for 20 min. After being washed twice, cells were analyzed with an Excalibur fluorescent activated cell sorter. The monoclonal antibodies 1C3, 1D12, 9C12, 20D4 and JA1, but not 8D3, were found to bind to a cell surface epitope of P503S.



In order to determine which tissues express P503S, immunohistochemical analysis was performed, essentially as described above, on a panel of normal tissues (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). HRP-labeled anti-mouse or anti-rabbit antibody followed by incubation with TMB was used to visualize P503S immunoreactivity. P503S was found to be highly expressed in prostate tissue, with lower levels of expression being observed in cervix, colon, ileum and kidney, and no expression being observed in adrenal, breast, duodenum, gall bladder, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis.

Western blot analysis was used to characterize anti-P503S monoclonal antibody specificity. SDS-PAGE was performed on recombinant (rec) P503S expressed in and purified from bacteria and on lysates from HEK293 cells transfected with full length P503S. Protein was transferred to nitrocellulose and then Western blotted with each of the anti-P503S monoclonal antibodies (20D4, JA1, 1D12, 6D12 and 9C12) at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to either a goat anti-mouse monoclonal antibody or to protein A-sepharose. The monoclonal antibody 20D4 detected the appropriate molecular weight 14 kDa recombinant P503S (amino acids 113-241) and the 23.5 kDa species in the HEK293 cell lysates transfected with full length P503S. Other anti-P503S monoclonal antibodies displayed similar specificity by Western blot.

#### **D) PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST P703P**

Rabbits were immunized with either a truncated (P703Ptr1; SEQ ID NO: 172) or full-length mature form (P703Pfl; SEQ ID NO: 523) of recombinant P703P protein was expressed in and purified from bacteria as described above. Affinity purified polyclonal antibody was generated using immunogen P703Pfl or P703Ptr1 attached to a solid support. Rabbit monoclonal antibodies were isolated using SLAM technology at Immgenics Pharmaceuticals. Table VII below lists both the polyclonal and monoclonal antibodies that were generated against P703P.



Table VII

Antibody	Immunogen	Species/type
Aff. Purif. P703P (truncated); #2594	P703Ptrl	Rabbit polyclonal
Aff. Purif. P703P (full length); #9245	P703Pfl	Rabbit polyclonal
2D4	P703Ptrl	Rabbit monoclonal
8H2	P703Ptrl	Rabbit monoclonal
7H8	P703Ptrl	Rabbit monoclonal

5                   The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 8H2, 7H8 and 2D4 were determined and are provided in SEQ ID NO: 506-508, respectively.

Epitope mapping studies were performed as described above. Monoclonal antibodies 2D4 and 7H8 were found to specifically bind to the peptides of SEQ ID NO: 10 509 (corresponding to amino acids 145-159 of SEQ ID NO: 172) and SEQ ID NO: 510 (corresponding to amino acids 11-25 of SEQ ID NO: 172), respectively. The polyclonal antibody 2594 was found to bind to the peptides of SEQ ID NO: 511-514, with the polyclonal antibody 9427 binding to the peptides of SEQ ID NO: 515-517.

The specificity of the anti-P703P antibodies was determined by Western 15 blot analysis as follows. SDS-PAGE was performed on (1) bacterially expressed recombinant antigen; (2) lysates of HEK293 cells and Ltk<sup>-/-</sup> cells either untransfected or transfected with a plasmid expressing full length P703P; and (3) supernatant isolated from these cell cultures. Protein was transferred to nitrocellulose and then Western blotted using the anti-P703P polyclonal antibody #2594 at an antibody concentration of 1 ug/ml. Protein 20 was detected using horse radish peroxidase (HRP) conjugated to an anti-rabbit antibody. A 35 kDa immunoreactive band could be observed with recombinant P703P. Recombinant P703P runs at a slightly higher molecular weight since it is epitope tagged. In lysates and supernatants from cells transfected with full length P703P, a 30 kDa band corresponding to P703P was observed. To assure specificity, lysates from HEK293 cells stably transfected



with a control plasmid were also tested and were negative for P703P expression. Other anti-P703P antibodies showed similar results.

Immunohistochemical studies were performed as described above, using anti-P703P monoclonal antibody. P703P was found to be expressed at high levels in normal prostate and prostate tumor tissue but was not detectable in all other tissues tested (breast tumor, lung tumor and normal kidney).

## EXAMPLE 19

### CHARACTERIZATION OF CELL SURFACE EXPRESSION AND

### 10 CHROMOSOME LOCALIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

This example describes studies demonstrating that the prostate-specific antigen P501S is expressed on the surface of cells, together with studies to determine the probable chromosomal location of P501S.

15 The protein P501S (SEQ ID NO: 113) is predicted to have 11 transmembrane domains. Based on the discovery that the epitope recognized by the anti-P501S monoclonal antibody 10E3-G4-D3 (described above in Example 17) is intracellular, it was predicted that following transmembrane determinants would allow the prediction of extracellular domains of P501S. Fig. 9 is a schematic representation of the P501S protein showing the predicted location of the transmembrane domains and the intracellular epitope described in Example 17. Underlined sequence represents the predicted transmembrane domains, bold sequence represents the predicted extracellular domains, and italicized sequence represents the predicted intracellular domains. Sequence that is both bold and underlined represents sequence employed to generate polyclonal rabbit serum. The location of the transmembrane domains was predicted using HHMTOP as described by Tusnady and Simon (Principles Governing Amino Acid Composition of Integral Membrane Proteins: Applications to Topology Prediction, *J. Mol. Biol.* 283:489-506, 1998).



Based on Fig. 9, the P501S domain flanked by the transmembrane domains corresponding to amino acids 274-295 and 323-342 is predicted to be extracellular. The peptide of SEQ ID NO: 518 corresponds to amino acids 306-320 of P501S and lies in the predicted extracellular domain. The peptide of SEQ ID NO: 519, which is identical to the peptide of SEQ ID NO: 518 with the exception of the substitution of the histidine with an asparagine, was synthesized as described above. A Cys-Gly was added to the C-terminus of the peptide to facilitate conjugation to the carrier protein. Cleavage of the peptide from the solid support was carried out using the following cleavage mixture: trifluoroacetic acid:ethanediol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for two hours, the peptide was precipitated in cold ether. The peptide pellet was then dissolved in 10% v/v acetic acid and lyophilized prior to purification by C18 reverse phase hplc. A gradient of 5-60% acetonitrile (containing 0.05% TFA) in water (containing 0.05% TFA) was used to elute the peptide. The purity of the peptide was verified by hplc and mass spectrometry, and was determined to be >95%. The purified peptide was used to generate rabbit polyclonal antisera as described above.

Surface expression of P501S was examined by FACS analysis. Cells were stained with the polyclonal anti-P501S peptide serum at 10 µg/ml, washed, incubated with a secondary FITC-conjugated goat anti-rabbit Ig antibody (ICN), washed and analyzed for FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. To demonstrate specificity in these assays, B-LCL transduced with an irrelevant antigen (P703P) or nontransduced were stained in parallel. For FACS analysis of prostate tumor cell lines, Lncap, PC-3 and DU-145 were utilized. Prostate tumor cell lines were dissociated from tissue culture plates using cell dissociation medium and stained as above. All samples were treated with propidium iodide (PI) prior to FACS analysis, and data was obtained from PI-excluding (*i.e.*, intact and non-permeabilized) cells. The rabbit polyclonal serum generated against the peptide of SEQ ID NO: 519 was shown to specifically recognize the surface of cells transduced to express P501S, demonstrating that the epitope recognized by the polyclonal serum is extracellular.



To determine biochemically if P501S is expressed on the cell surface, peripheral membranes from Lncap cells were isolated and subjected to Western blot analysis. Specifically, Lncap cells were lysed using a dounce homogenizer in 5 ml of homogenization buffer (250 mM sucrose, 10 mM HEPES, 1mM EDTA, pH 8.0, 1  
 5 complete protease inhibitor tablet (Boehringer Mannheim)). Lysate samples were spun at 1000 g for 5 min at 4 °C. The supernatant was then spun at 8000g for 10 min at 4 °C. Supernatant from the 8000g spin was recovered and subjected to a 100,000g spin for 30 min at 4 °C to recover peripheral membrane. Samples were then separated by SDS-PAGE and Western blotted with the mouse monoclonal antibody 10E3-G4-D3 (described above in  
 10 Example 17) using conditions described above. Recombinant purified P501S, as well as HEK293 cells transfected with and over-expressing P501S were included as positive controls for P501S detection. LCL cell lysate was included as a negative control. P501S could be detected in Lncap total cell lysate, the 8000g (internal membrane) fraction and also in the 100,000g (plasma membrane) fraction. These results indicate that P501S is  
 15 expressed at, and localizes to, the peripheral membrane.

To demonstrate that the rabbit polyclonal antiserum generated to the peptide of SEQ ID NO: 519 specifically recognizes this peptide as well as the corresponding native peptide of SEQ ID NO: 518, ELISA analyses were performed. For these analyses, flat-bottomed 96 well microtiter plates were coated with either the peptide of SEQ ID NO: 519,  
 20 the longer peptide of SEQ ID NO: 520 that spans the entire predicted extracellular domain, the peptide of SEQ ID NO: 521 which represents the epitope recognized by the P501S-specific antibody 10E3-G4-D3, or a P501S fragment (corresponding to amino acids 355-526 of SEQ ID NO: 113) that does not include the immunizing peptide sequence, at 1 µg/ml for 2 hours at 37 °C. Wells were aspirated, blocked with phosphate buffered saline  
 25 containing 1% (w/v) BSA for 2 hours at room temperature and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified anti-P501S polyclonal rabbit serum was added at 2 fold dilutions (1000 ng - 125 ng) in PBST and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and incubating with HRP-conjugated goat anti-rabbit IgG (H+L) Affinipure F(ab') fragment at 1:20000 for 30 min.



Plates were then washed and incubated for 15 min in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 11, the anti-P501S polyclonal rabbit serum specifically recognized the peptide of SEQ ID NO: 519 used in the immunization as well as the longer peptide of SEQ ID NO: 520, but did not recognize the irrelevant P501S-derived peptides and fragments.

In further studies, rabbits were immunized with peptides derived from the P501S sequence and predicted to be either extracellular or intracellular, as shown in Fig. 9. Polyclonal rabbit sera were isolated and polyclonal antibodies in the serum were purified, as described above. To determine specific reactivity with P501S, FACS analysis was employed, utilizing either B-LCL transduced with P501S or the irrelevant antigen P703P, of B-LCL infected with vaccinia virus-expressing P501S. For surface expression, dead and non-intact cells were excluded from the analysis as described above. For intracellular staining, cells were fixed and permeabilized as described above. Rabbit polyclonal serum generated against the peptide of SEQ ID NO: 548, which corresponds to amino acids 181-198 of P501S, was found to recognize a surface epitope of P501S. Rabbit polyclonal serum generated against the peptide SEQ ID NO: 551, which corresponds to amino acids 543-553 of P501S, was found to recognize an epitope that was either potentially extracellular or intracellular since in different experiments intact or permeabilized cells were recognized by the polyclonal sera. Based on similar deductive reasoning, the sequences of SEQ ID NO: 541-547, 549 and 550, which correspond to amino acids 109-122, 539-553, 509-520, 37-54, 342-359, 295-323, 217-274, 143-160 and 75-88, respectively, of P501S, can be considered to be potential surface epitopes of P501S recognized by antibodies.

The chromosomal location of P501S was determined using the GeneBridge 4 Radiation Hybrid panel (Research Genetics). The PCR primers of SEQ ID NO: 528 and 529 were employed in PCR with DNA pools from the hybrid panel according to the manufacturer's directions. After 38 cycles of amplification, the reaction products were separated on a 1.2% agarose gel, and the results were analyzed through the Whitehead



Institute/MIT Center for Genome Research web server (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) to determine the probable chromosomal location. Using this approach, P501S was mapped to the long arm of chromosome 1 at WI-9641 between q32 and q42. This region of chromosome 1 has been linked to prostate cancer susceptibility in hereditary prostate cancer (Smith *et al. Science* 274:1371-1374, 1996 and Berthon *et al. Am. J. Hum. Genet.* 62:1416-1424, 1998). These results suggest that P501S may play a role in prostate cancer malignancy.

## EXAMPLE 20

### 10 REGULATION OF EXPRESSION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

Steroid (androgen) hormone modulation is a common treatment modality in prostate cancer. The expression of a number of prostate tissue-specific antigens have previously been demonstrated to respond to androgen. The responsiveness of the prostate-specific antigen P501S to androgen treatment was examined in a tissue culture system as follows.

Cells from the prostate tumor cell line LNCaP were plated at  $1.5 \times 10^6$  cells/T75 flask (for RNA isolation) or  $3 \times 10^5$  cells/well of a 6-well plate (for FACS analysis) and grown overnight in RPMI 1640 media containing 10% charcoal-stripped fetal calf serum (BRL Life Technologies, Gaithersburg, MD). Cell culture was continued for an additional 72 hours in RPMI 1640 media containing 10% charcoal-stripped fetal calf serum, with 1 nM of the synthetic androgen Methyltrienolone (R1881; New England Nuclear) added at various time points. Cells were then harvested for RNA isolation and FACS analysis at 0, 1, 2, 4, 8, 16, 24, 28 and 72-hours post androgen addition. FACS analysis was performed using the anti-P501S antibody 10E3-G4-D3 and permeabilized cells.

For Northern analysis, 5-10 micrograms of total RNA was run on a formaldehyde denaturing gel, transferred to Hybond-N nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ), cross-linked and stained with methylene blue. The



filter was then prehybridized with Church's Buffer (250 mM Na<sub>2</sub>HPO<sub>4</sub>, 70 mM H<sub>3</sub>PO<sub>4</sub>, 1 mM EDTA, 1% SDS, 1% BSA in pH 7.2) at 65 °C for 1 hour. P501S DNA was labeled with 32P using High Prime random-primed DNA labeling kit (Boehringer Mannheim). Unincorporated label was removed using MicroSpin S300-HR columns (Amersham Pharmacia Biotech). The RNA filter was then hybridized with fresh Church's Buffer containing labeled cDNA overnight, washed with 1X SCP (0.1 M NaCl, 0.03 M Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O, 0.001 M Na<sub>2</sub>EDTA), 1% sarkosyl (n-lauroylsarcosine) and exposed to X-ray film.

Using both FACS and Northern analysis, P501S message and protein levels were found in increase in response to androgen treatment.

## EXAMPLE 20

### PREPARATION OF FUSION PROTEINS OF PROSTATE-SPECIFIC ANTIGENS

The example describes the preparation of a fusion protein of the prostate-specific antigen P703P and a truncated form of the known prostate antigen PSA. The truncated form of PSA has a 21 amino acid deletion around the active serine site. The expression construct for the fusion protein also has a restriction site at 3' end, immediately prior to the termination codon, to aid in adding cDNA for additional antigens.

The full-length cDNA for PSA was obtained by RT-PCR from a pool of RNA from human prostate tumor tissues using the primers of SEQ ID NO: 607 and 608, and cloned in the vector pCR-Blunt II-TOPO. The resulting cDNA was employed as a template to make two different fragments of PSA by PCR with two sets of primers (SEQ ID NO: 609 and 610; and SEQ ID NO: 611 and 612). The PCR products having the expected size were used as templates to make truncated forms of PSA by PCR with the primers of SEQ ID NO: 611 and 613, which generated PSA (delta 208-218 in amino acids). The cDNA for the mature form of P703P with a 6X histidine tag at the 5' end, was prepared by PCR with P703P and the primers of SEQ ID NO: 614 and 615. The cDNA for the fusion of P703P with the truncated form of PSA (referred to as FOPP) was then



obtained by PCR using the modified P703P cDNA and the truncated form of PSA cDNA as templates and the primers of SEQ ID NO: 614 and 615. The FOPP cDNA was cloned into the NdeI site and XhoI site of the expression vector pCRX1, and confirmed by DNA sequencing. The determined cDNA sequence for the fusion construct FOPP is provided in  
 5 SEQ ID NO: 616, with the amino acid sequence being provided in SEQ ID NO: 617.

## EXAMPLE 21

### REAL-TIME PCR CHARACTERIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S IN PERIPHERAL BLOOD OF PROSTATE CANCER PATIENTS

10

Circulating epithelial cells were isolated from fresh blood of normal individuals and metastatic prostate cancer patients, mRNA isolated and cDNA prepared using real-time PCR procedures. Real-time PCR was performed with the Taqman<sup>TM</sup> procedure using both gene specific primers and probes to determine the levels of gene  
 15 expression.

Epithelial cells were enriched from blood samples using an immunomagnetic bead separation method (Dynal A.S., Oslo, Norway). Isolated cells were lysed and the magnetic beads removed. The lysate was then processed for poly A<sup>+</sup> mRNA isolation using magnetic beads coated with Oligo(dT)25. After washing the beads in  
 20 buffer, bead/poly A<sup>+</sup> RNA samples were suspended in 10 mM Tris HCl pH 8.0 and subjected to reversed transcription. The resulting cDNA was subjected to real-time PCR using gene specific primers. Beta-actin content was also determined and used for normalization. Samples with P501S copies greater than the mean of the normal samples + 3 standard deviations were considered positive. Real time PCR on blood samples was  
 25 performed using the Taqman<sup>TM</sup> procedure but extending to 50 cycles using forward and reverse primers and probes specific for P501S. Of the eight samples tested, 6 were positive for P501S and  $\beta$ -actin signal. The remaining 2 samples had no detectable  $\beta$ -actin or P501S. No P501S signal was observed in the four normal blood samples tested.







## CLAIMS

What is claimed:

1. An isolated polypeptide, comprising at least an immunogenic portion of a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779 under moderately stringent conditions; and

(c) complements of sequences of (a) or (b).

2. An isolated polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434,



435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779, or a complement of any of the foregoing polynucleotide sequences.

3. An isolated polypeptide comprising a sequence recited in any one of SEQ ID NO: 108, 112, 113, 114, 172, 176, 178, 327, 329, 331, 338, 339, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 780, 781, 810, 811 and 814.

4. An isolated polynucleotide encoding at least 15 amino acid residues of a prostate-specific protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779, or a complement of any of the foregoing sequences.

5. An isolated polynucleotide encoding a prostate-specific protein, or a variant thereof, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOs: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-



461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779, or a complement of any of the foregoing sequences.

6. An isolated polynucleotide, comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779.

7. An isolated polynucleotide, comprising a sequence that hybridizes to a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779 under moderately stringent conditions.

8. An isolated polynucleotide complementary to a polynucleotide according to any one of claims 4-7.

9. An expression vector, comprising a polynucleotide according to any one of claims 4-8.

10. A host cell transformed or transfected with an expression vector according to claim 9.



11. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a prostate-specific protein that comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779, or a complement of any of the foregoing polynucleotide sequences.

12. A fusion protein, comprising at least one polypeptide according to claim 1.

13. A fusion protein according to claim 12, wherein the fusion protein comprises an expression enhancer that increases expression of the fusion protein in a host cell transfected with a polynucleotide encoding the fusion protein.

14. A fusion protein according to claim 12, wherein the fusion protein comprises a T helper epitope that is not present within the polypeptide of claim 1.

15. A fusion protein according to claim 12, wherein the fusion protein comprises an affinity tag.

16. An isolated polynucleotide encoding a fusion protein according to claim 12.

17. A pharmaceutical composition, comprising a physiologically acceptable carrier and at least one component selected from the group consisting of:

(a) a polypeptide according to claim 1;



- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

18. An immunogenic composition comprising an immunostimulant and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;
- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

19. An immunogenic composition according to claim 18, wherein the immunostimulant is an adjuvant.

20. An immunogenic composition according to claim 18, wherein the immunostimulant induces a predominantly Type I response.

21. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a pharmaceutical composition according to claim 17.

22. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an immunogenic composition according to claim 18.



23. A pharmaceutical composition comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with a pharmaceutically acceptable carrier or excipient.

24. A pharmaceutical composition according to claim 23, wherein the antigen presenting cell is a dendritic cell or a macrophage.

25. An immunogenic composition comprising an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii);  
in combination with an immunostimulant.

26. An immunogenic composition according to claim 25, wherein the immunostimulant is an adjuvant.

27. An immunogenic composition according to claim 25, wherein the immunostimulant induces a predominantly Type I response.

28. An immunogenic composition according to claim 25, wherein the antigen-presenting cell is a dendritic cell.



29. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii) encoded by a polynucleotide recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789;

and thereby inhibiting the development of a cancer in the patient.

30. A method according to claim 29, wherein the antigen-presenting cell is a dendritic cell.

31. A method according to any one of claims 21, 22 and 29, wherein the cancer is prostate cancer.

32. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:



(i) polynucleotides recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789; and

(ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

33. A method according to claim 32, wherein the biological sample is blood or a fraction thereof.

34. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 32.

35. A method for stimulating and/or expanding T cells specific for a prostate-specific protein, comprising contacting T cells with at least one component selected from the group consisting of:

(a) polypeptides comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789;

(ii) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 under moderately stringent conditions; and

(iii) complements of sequences of (i) or (ii);



- (b) polynucleotides encoding a polypeptide of (a); and
- (c) antigen presenting cells that express a polypeptide of (a);

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

36. An isolated T cell population, comprising T cells prepared according to the method of claim 35.

37. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population according to claim 36.

38. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789;

(2) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 under moderately stringent conditions; and

(3) complements of sequences of (1) or (2);

(ii) polynucleotides encoding a polypeptide of (i); and

(iii) antigen presenting cells that expresses a polypeptide of (i);



such that T cells proliferate; and

(b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient.

39. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789;

(2) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 under moderately stringent conditions; and

(3) complements of sequences of (1) or (2);

(ii) polynucleotides encoding a polypeptide of (i); and

(iii) antigen presenting cells that express a polypeptide of (i);

such that T cells proliferate;

(b) cloning at least one proliferated cell to provide cloned T cells; and

(c) administering to the patient an effective amount of the cloned T cells, and thereby inhibiting the development of a cancer in the patient.

40. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:



(a) contacting a biological sample obtained from a patient with a binding agent that binds to a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

41. A method according to claim 40, wherein the binding agent is an antibody.

42. A method according to claim 43, wherein the antibody is a monoclonal antibody.

43. A method according to claim 40, wherein the cancer is prostate cancer.

44. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent;



(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

45. A method according to claim 44, wherein the binding agent is an antibody.

46. A method according to claim 45, wherein the antibody is a monoclonal antibody.

47. A method according to claim 44, wherein the cancer is a prostate cancer.

48. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789, or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

49. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.



50. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

51. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777 and 789, or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

52. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

53. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

54. A diagnostic kit, comprising:

(a) one or more antibodies according to claim 11; and

(b) a detection reagent comprising a reporter group.



55. A kit according to claim 54, wherein the antibodies are immobilized on a solid support.

56. A kit according to claim 54, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

57. A kit according to claim 54, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

58. An oligonucleotide comprising 10 to 40 contiguous nucleotides that hybridize under moderately stringent conditions to a polynucleotide that encodes a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779, or a complement of any of the foregoing polynucleotides.

59. A oligonucleotide according to claim 58, wherein the oligonucleotide comprises 10-40 contiguous nucleotides recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772 and 779.



60. A diagnostic kit, comprising:
- (a) an oligonucleotide according to claim 59; and
  - (b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.



COMPOSITIONS AND METHODS FOR THE THERAPY  
AND DIAGNOSIS OF PROSTATE CANCER

ABSTRACT OF THE DISCLOSURE

Compositions and methods for the therapy and diagnosis of cancer, such as prostate cancer, are disclosed. Compositions may comprise one or more prostate-specific proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a prostate-specific protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as prostate cancer. Diagnostic methods based on detecting a prostate-specific protein, or mRNA encoding such a protein, in a sample are also provided.



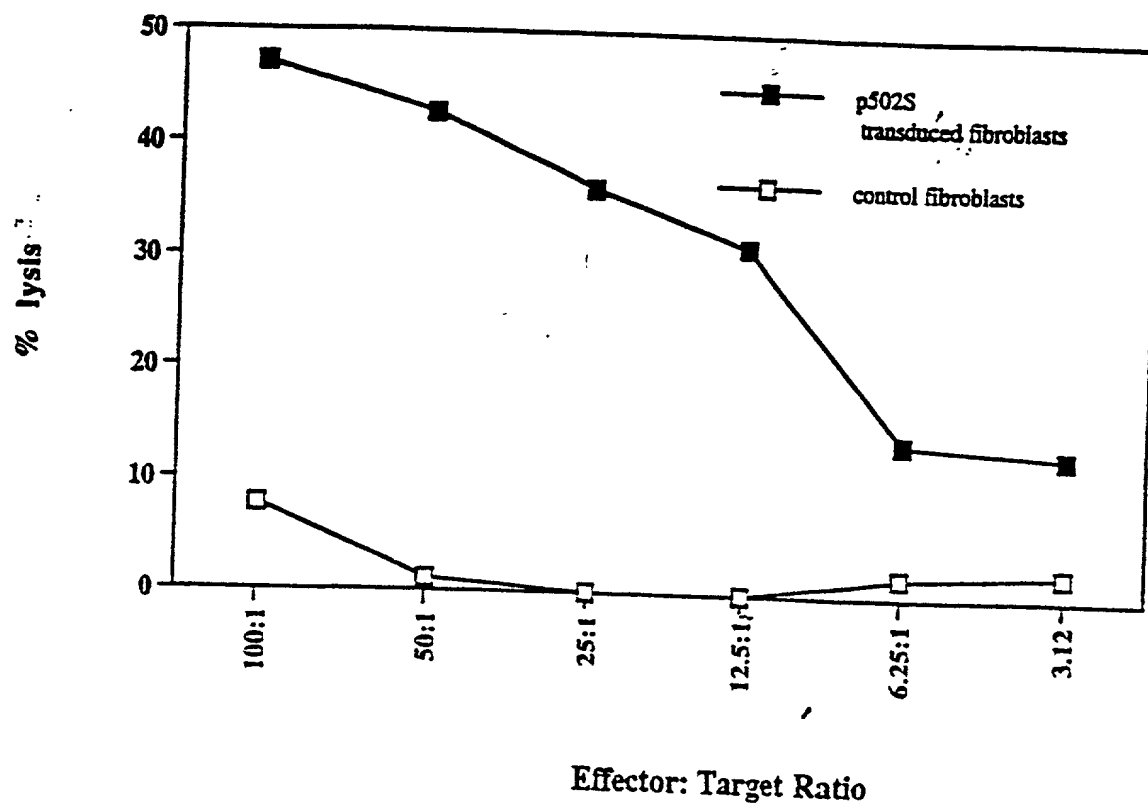


FIG. 1



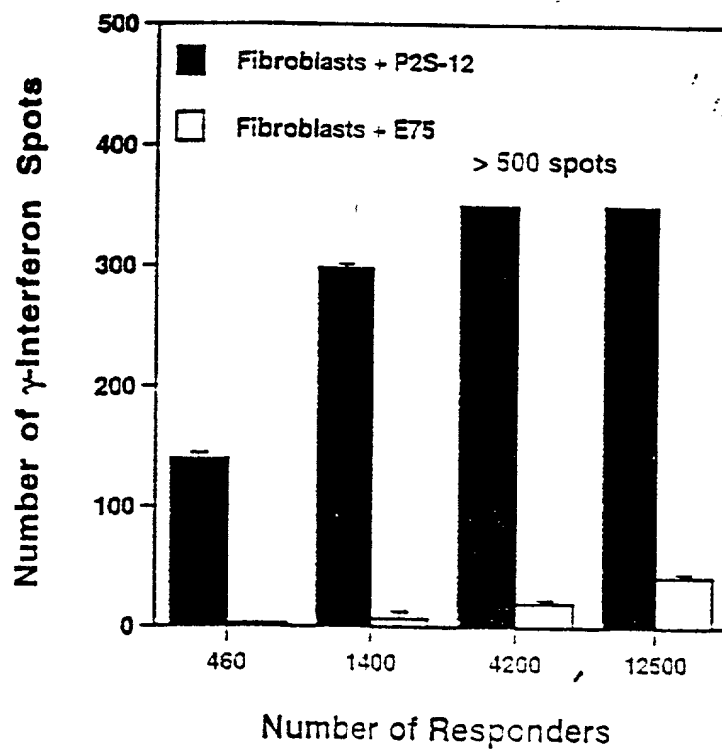
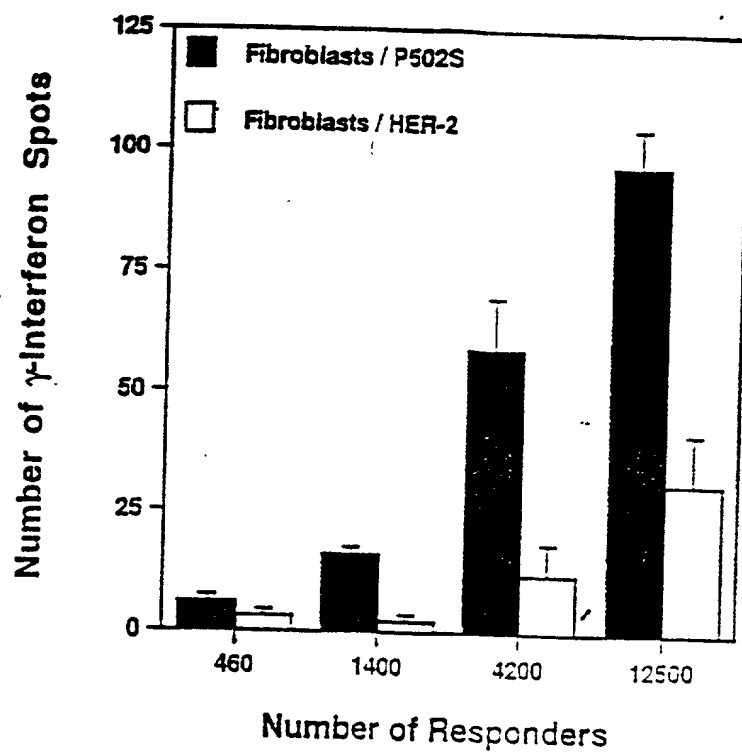


FIG. 2A

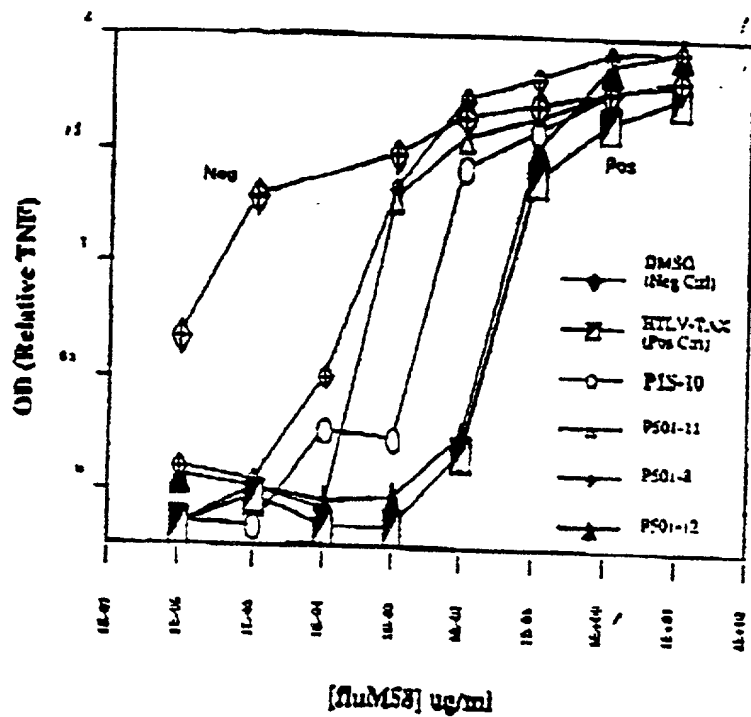




*FIG. 2B*



002790-66/66560



Figure

3



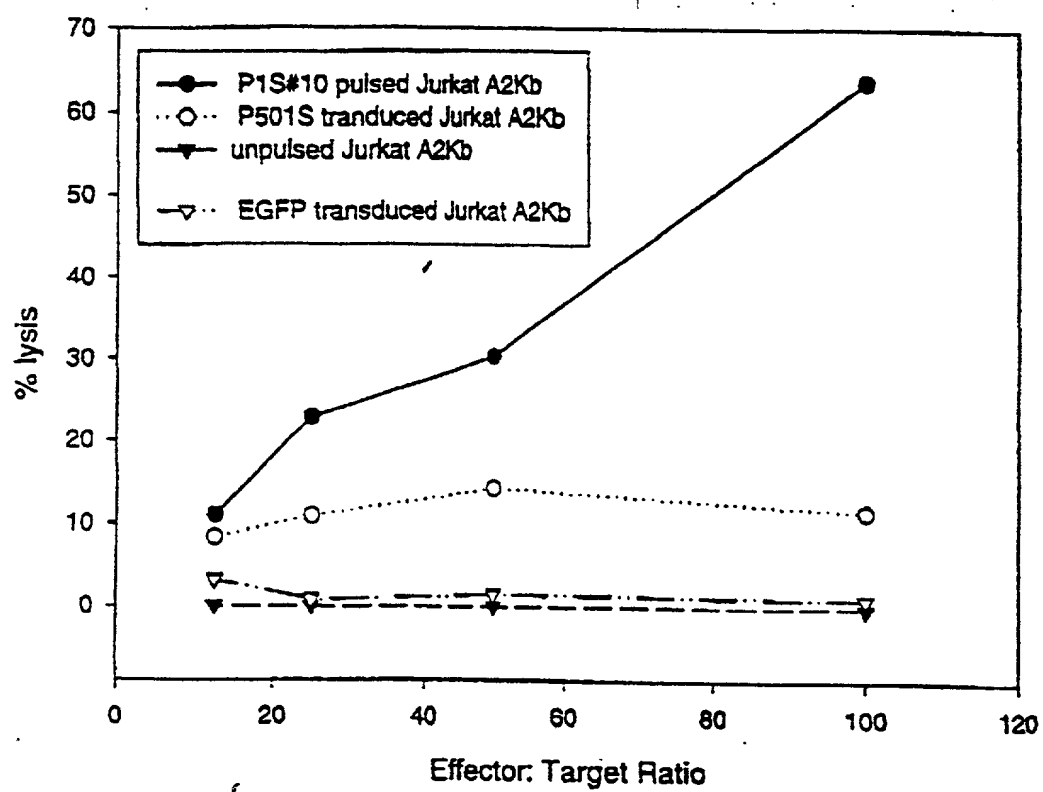


Figure 4



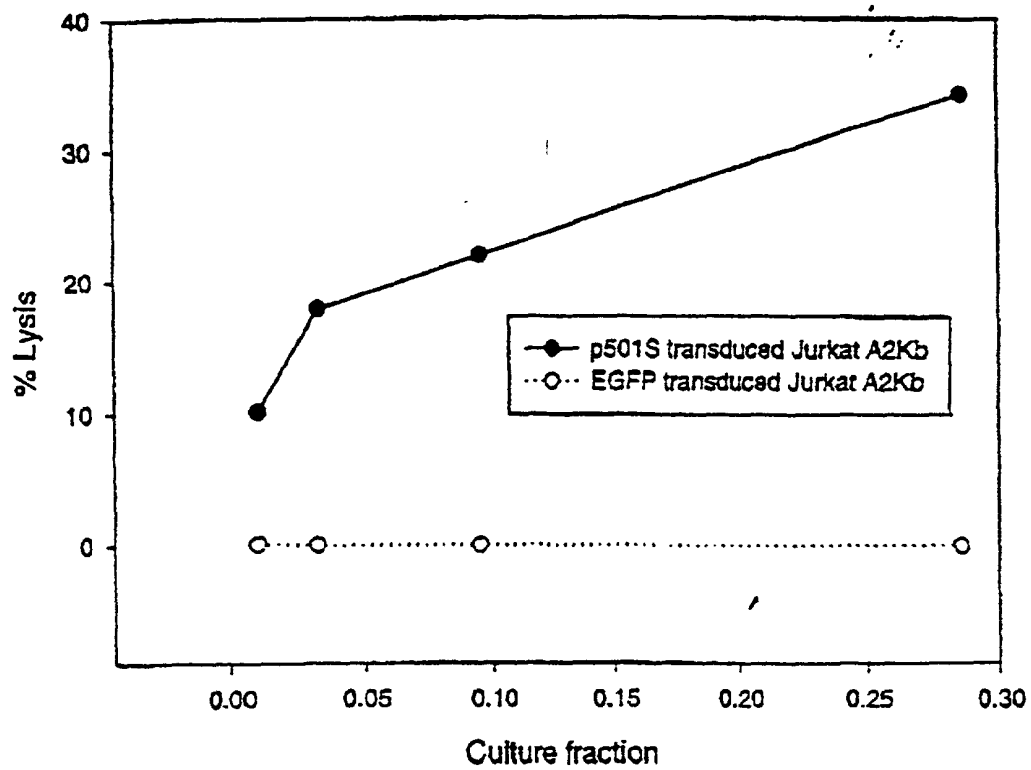


Figure 5



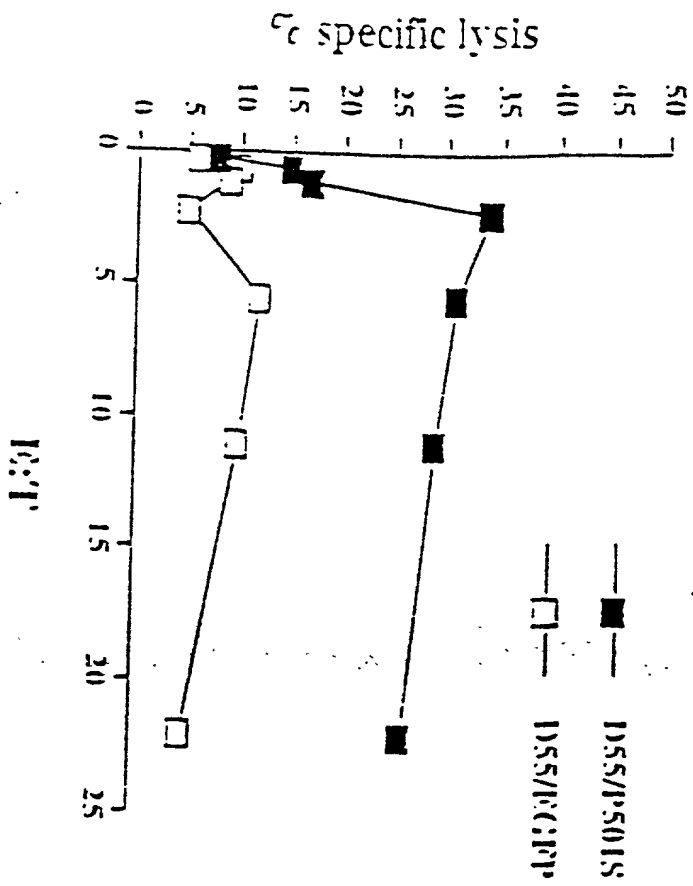


Fig. 6A

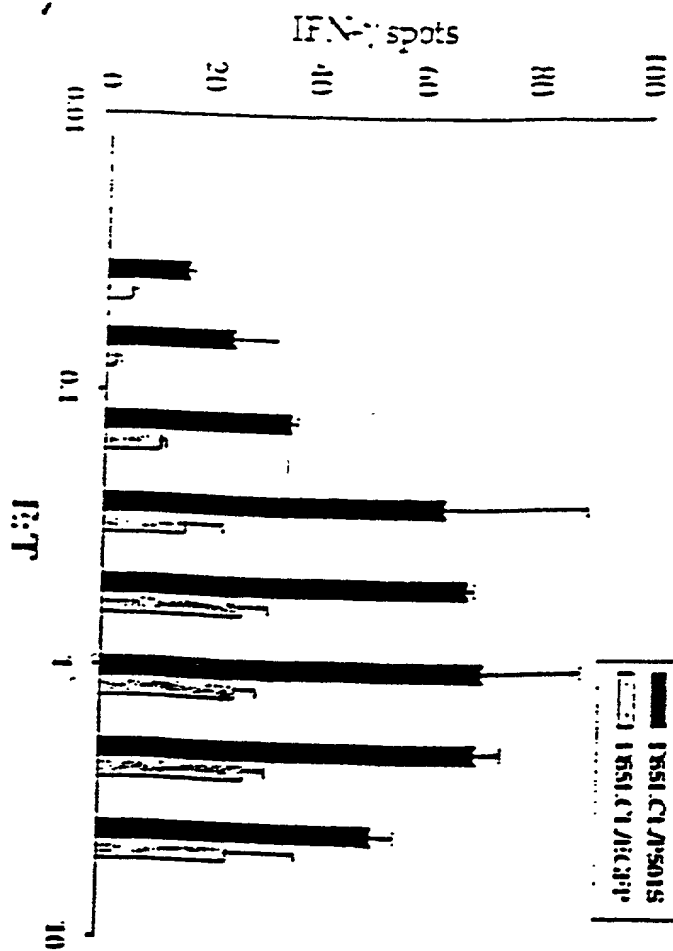
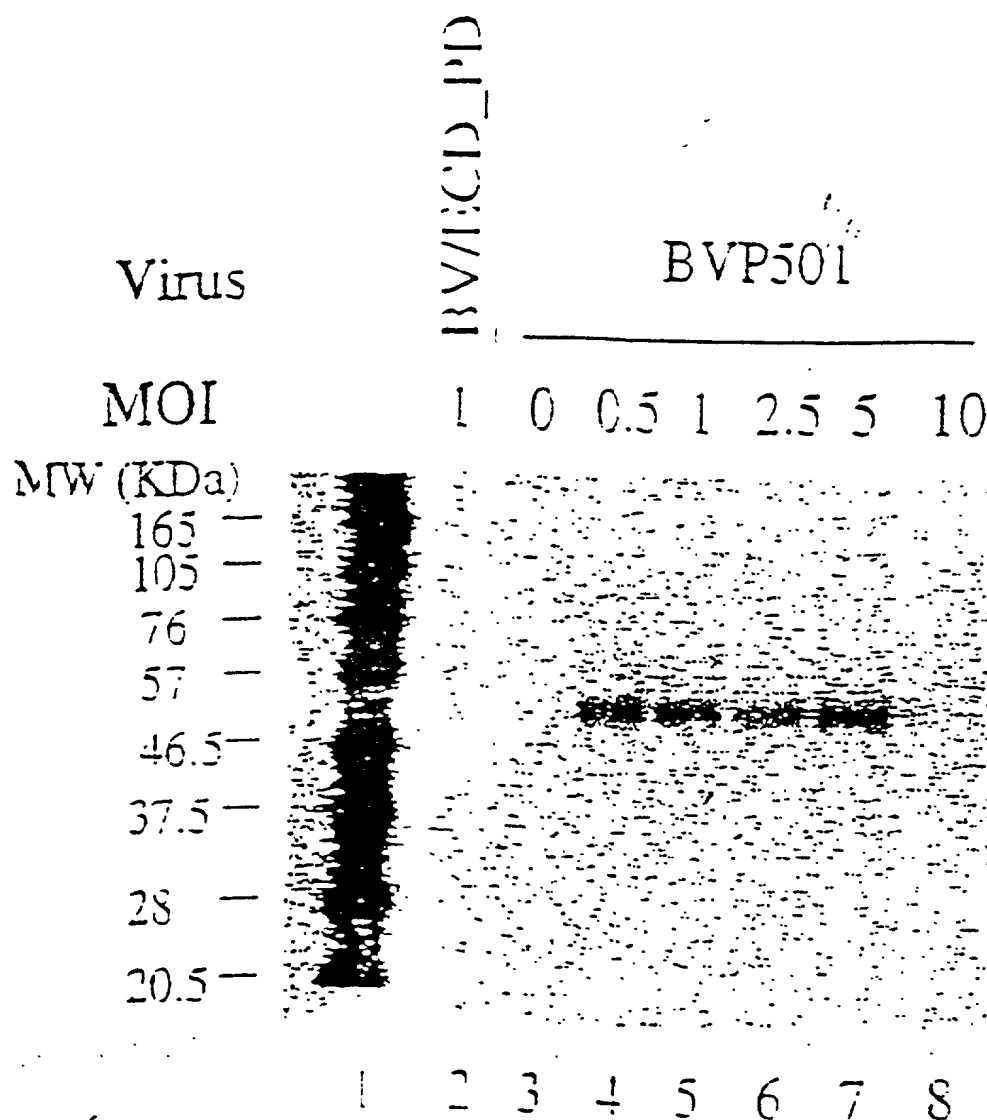


Fig. 6B



# Expression of P501S by the Baculovirus Expression System



0.6 million high titer cells in 6-well plate were infected with an unrelated control virus BV/ECD\_PD (lane 2), without virus (lane 3), or with recombinant baculovirus for P501 at different MOI (lane 4 - 8). Cell lysates were run on SDS-PAGE under the reducing condition and analysed by Western blot with a monoclonal antibody against P501S-10E9-G4D3. Lane 1 is the biotinylated protein molecular weight marker (KDa).

Fig. 7







7

# Figure 1. Schematic of P501S with predicted transmembrane, cytoplasmic, and extracellular regions

MVQRLVYSRLIRK AQLALYNLLTTCILEYCTLAAGT YVPLLEKVGVERKFM TMVLGIGPVILGLCYPLLDGASAS  
 DWWRGRYGRRRP EIWALSLGILLSEFLPRAGWL AGLLCTDDRRPLE LALLHGVCLLDGQGVCTPL  
 LALLSLERDPDHCQ AYSYVAFMISLGGCTGYTPAI DWDTSATAPVLCQDRE  
 CLPGLLEFLPCTCYAATLLY AEFALGPTEFAEGLSAPSSPHTCPCHRAFAFRLGAILPRI  
 HQLCCRPPTLR LLYAFLEQSWMAI NLETTFTYTP YGEGLYGCVPRAMPCTEARHHYDEGVH  
 MGSILQLFLOCAISLYFSLYM DRIVQRFCTRAVYLAS VAAFPYAAATCLSHSYAVVTA SAA  
 LTGETFSALQILPYTLASLY HREKQVFLPKYRGDTGCGASSEDSTATSEFLCPKPGAFPNGHIVGAGGSQL  
 LPPPPALCGASACDVSVRVVGEPTFAKVVPERG ICLDLALHDSAFLLSQVAPSLF MGSIVQLSQS  
 VTAYMVSAAGILGLVAIFYAT QVVFDKSLAKTSA

Underlined sequence: Predicted transmembrane domain; Bold sequence: Predicted extracellular domain;  
 Italic sequence: Predicted intracellular domain. Sequence in bold/underlined: used to generate polyclonal rabbit serum

Localization of domains predicted using IMMTOPI (C.P. Tusnady and I. Simon (1998) Principles  
 Governing Amino Acid Composition of Integral Membrane Proteins: Applications to topology Prediction. J.Mol Biol. 283,  
 489-506.



# Genomic Map of (5) Corlxa Candidate Genes

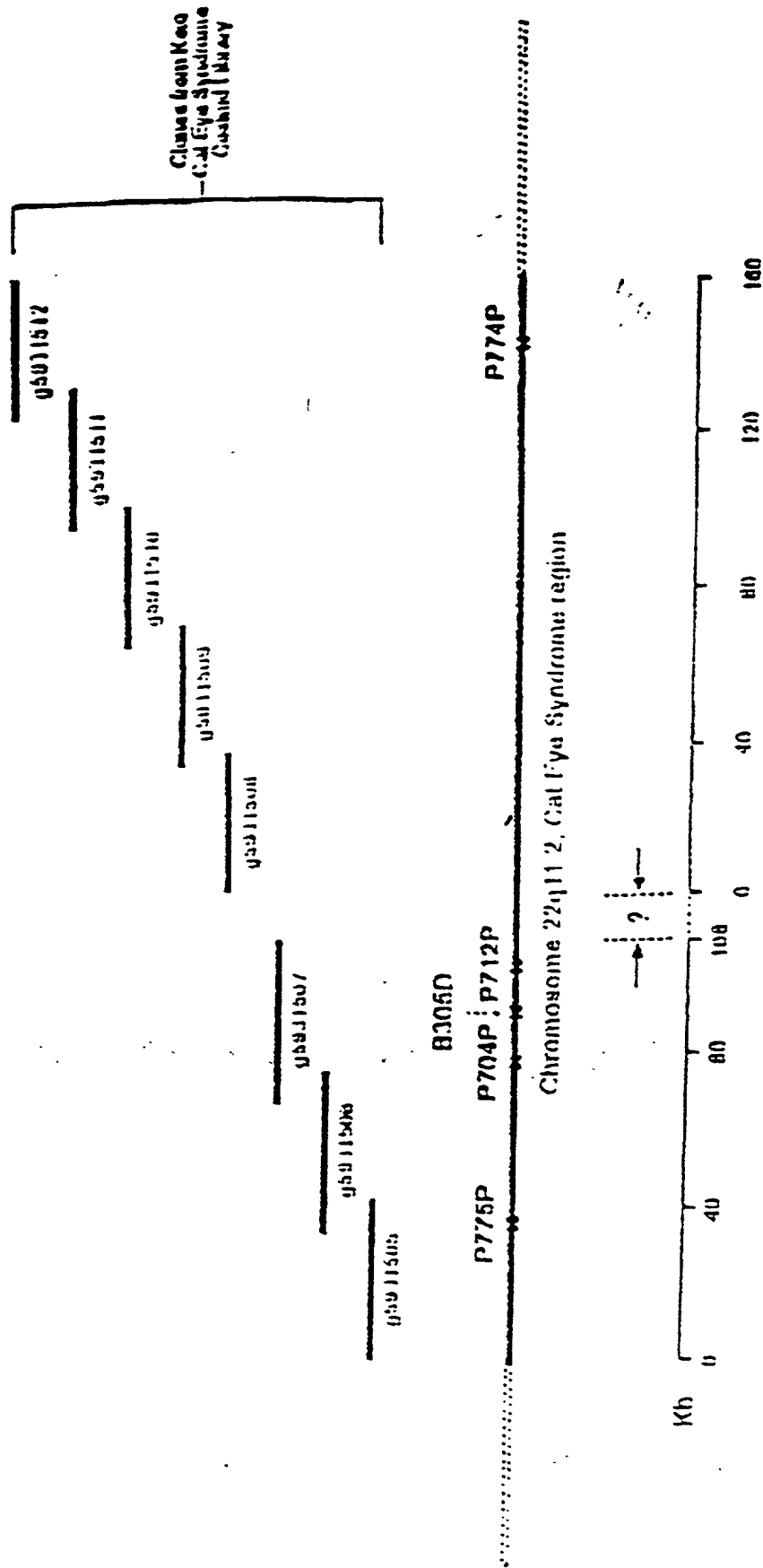


Fig. 10



# FIGURE 4. ELISA assay of rabbit polyclonal antibody specificity

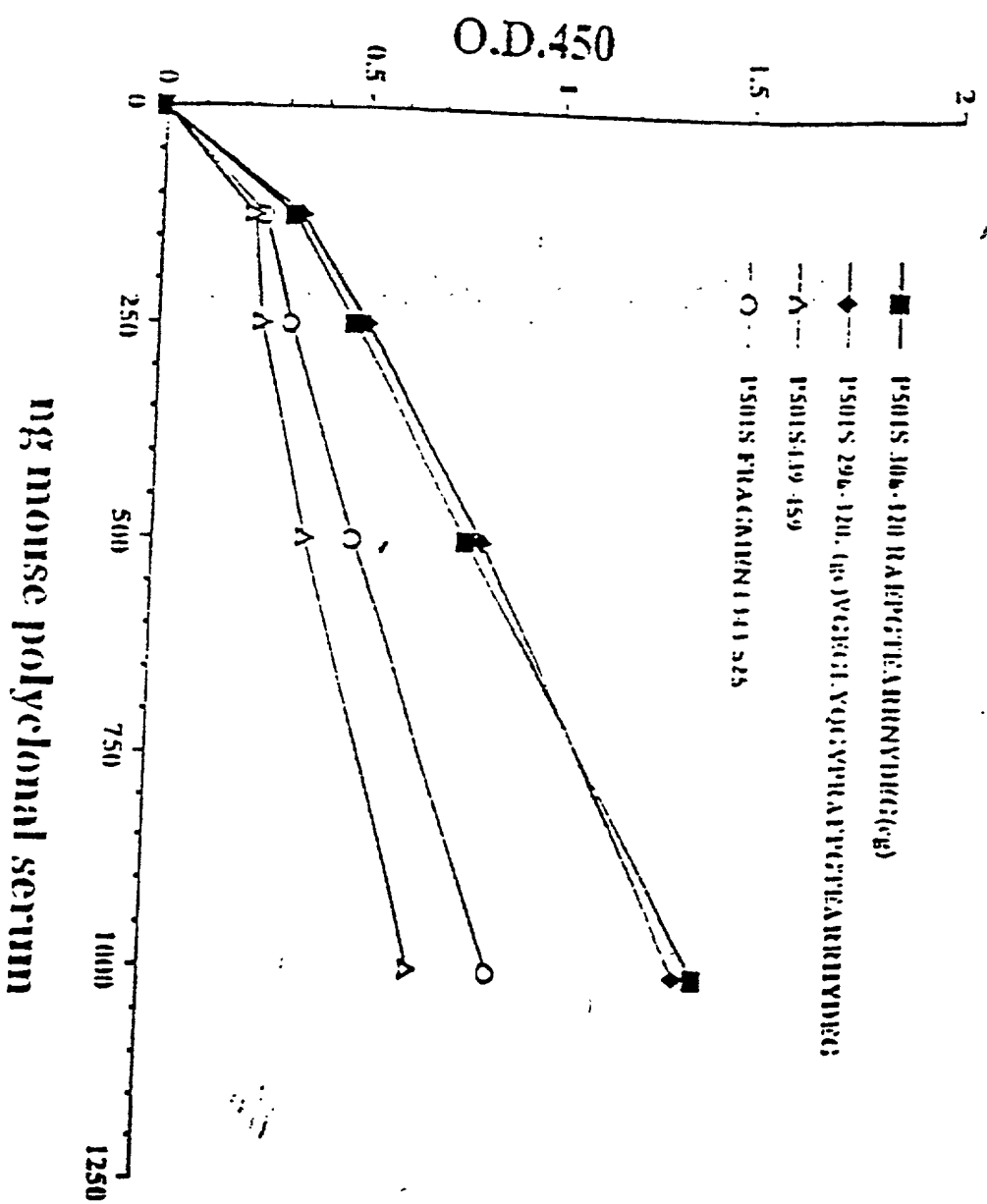


Fig. 11



10 20 30 40 50 60 70  
 GTCACCTTAGGAAAAGGTGTCCTTTCTGGGCAGCCGGGCTCAGCATGAGGAACAGAAGGAATGACACTCTGG 70  
 ACAGCACCCCGGACCCTGTACTCCAGCGCGTCTCGGAGCACAGACTTGTCTTACACTGAAAAGCGACTTGGT 140  
 GAATTTTATTCAAGCAAATTTTAAGAAACGAGAATGTGTCTTCTTTACCAAAGATTCCAAGGCCACGGAG 210  
 AATGTGTGCAAGTGTGGCTATGCCCGAGAGCCAGCAGATGGAAGGCACCCAGATCAACCAAAGTGAGAAAT 280  
 GGAACTACAAGAAACACACCAAGGAATTTCTTACCAGCCCTTTGGGGATATTTCAGTTTGAGACACTGGG 350  
 360 370 380 390 400 410 420  
 GAAGAAAAGGGGAAGTATATACGTCTGTCTTGCACACGGACCGGAAATCCTTTACGAGCTGCTGACCCAG 420  
 CACTGGCACCTGAAAACAACCAACCTGGTCATTTCTGTGACCGGGGGGGCCCAAGAACTTCGCCCTGAAGC 490  
 CGCGCATGCGCAAGATCTTCAGCCGGCTCATCTACATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGG 560  
 AGGCACCCATTATGGCCTGACGAAGTACATCGGGGAGGTGGTGAGAGATAACACCATCAGCAGGAGTTCA 630  
 GAGGAGAATATTGTGGCCATTGGCATAGCAGCTTGGGGCATGGTCTCCAAACCGGGACACCCCTCATCAGGA 700  
 710 720 730 740 750 760 770  
 ATTGGGATGCTGAGGGCTATTTTTTAGCCCAAGTACCTTATGGATGACTTCACAAGGGATCCACTGTATAT 770  
 CCTGGACAACAACCAACACACATTTGGCTGCTGGTGGACAAATGGCTGTGATGGACATCCCACTGTGGAAGCA 840  
 AAGCTCCGGGAATCAGCTAGAGAAAGCATATCTGTGAGCGCACTATTCAAGATTCCAACTATGGTGGCAAGA 910  
 TCCCCATTGTGTGTTTTGCCCCAAGGAGGTGGAAAAGAGACTTTGAAAGCCATCAATAGCTCCATCAAAAA 980  
 TAAAAATTCCTTGTTGGTGGTGGAAAGGCTCGGGCGGATCGCTGATGTGATCGCTAGCCTGGTGGAGGTG 1050  
 1060 1070 1080 1090 1100 1110 1120  
 GAGGATGCCCGGACATCTTTCTCCCGTCAAGGAGAAGCTGGTGGCTTTTTTACCCCGGACGGTGTCTCGGG 1120  
 TGTCTGAGGAGGAGACTGAGAGTTGGATCAAAATGGCTCAAAAGAAATTTCTGAAATGTTCTCACCTATTAAC 1190  
 AGTTATTAAATGGAAGAAGCTGGGGATGAAATTTGTAGCAATGCCATCTCTACGCTCTATACAAAAGCC 1260  
 TTCAGCACCAAGTGAGCAAGACAAAGGATAACTGGAATGGGC-GTGAAGCTTCTGCTGGAGTGAATCCAGC 1330  
 TGGACTTAGCCCAATGATGAGATTTTACCAATGACCGCCGATGGGAGTCTGCTGACCTTCAAGAAGTCAAT 1400  
 1410 1420 1430 1440 1450 1460 1470  
 GTTTACGGCTCTCATAAAGGACAGACCCAAGTTGTCTGGCTCTTTCTGGAGAATGGCTTGAACCTACGG 1470  
 AAGTTTCTCACCCATGATGTCTCACTGAACCTCTCTCCAACCACTTCAGCACGCTTGTGTACCGGAATC 1540  
 TGCAGATCGCCAAGAATTTCTATAATGATGCCCTCTCTACGTTTGTCTGGAACTGGTTGCGAATTTCCG 1610  
 AAGAGGCTTCCGGAAGGAAGACAGAAATGGCGGGGAGAGATGGACATAGAACTCCACGACGTGTCTCT 1680  
 ATTACTCGGCACCCCTGCAAGCTCTCTTCATCTGGGCCATTCTTCAGAAAGGAAGGAATCTTCCAAAG 1750  
 1760 1770 1780 1790 1800 1810 1820  
 TCATTTGGGAGCAGAACCAAGGGCTGCACTCTGGCAGCCCTGCGAGCCAGCAAGCTTCTGAAGACTCTGGC 1820  
 CAAAGTGAAGAAAGACATCAATGCTGTGTGGGAGTTCGAGGAGCTGGCTAATGAGTACGAGACCTCGGGCT 1890  
 GTTGAGCTGTCACTGAGTGTACAGCAGCGATGAAGACTTGGCAGAAACAGCTGCTGGTGTATTCTGTG 1960  
 AAGCTTGGGGTGGAAAGCAACTGTCTGGAGCTGGGGTGGAGGCAACAGACCAAGCATTCACCGCCCAAGC 2030  
 TGGGGTCCAGAAATTTCTTTCTAAGCAATGGATGGAGAGATTTCCCGAGACACCAAGAACTGGAAGATT 2100

Fig. 12A (i)



00503703-061300

2110	2120	2130	2140	2150	2160	2170
TCCTGTGTCTGTTTATATACCTTGGTGGGCTGTGGCTTTGTATCATTTAGGAAGAAACCTGTGACA	2170					
AGCACAAGAAGCTGCTTTGGTACTATGTGGGCTCTTCACCTCCCCCTTCGTGGTCTTCTCCTGGAATGT	2240					
GGTCTTCTACATCGCCTTCTCCTGCTGTTTGGCTACGTGCTGCTCATGGATTTCCATTGGGTGCCACAC	2310					
CCCCCGAGCTGCTCCTGTACTCCCTGGTCTTTGCTCTTCTGTGATGAAGTCAGACAGTGGTACGTAA	2380					
ATGGGGTGAATTATTTTACTGACCTGTGGAATGTGATGGACACGCTGGGGCTTTTTTACTTCATAGCAGG	2450					
2460	2470	2480	2490	2500	2510	2520
AATTGTATTTGGGCTCCACTCTTCTAATAAAAGCTCTTTGTATTCTGGACGAGTCATTTTCTGTCTGGAC	2520					
TACATTATTTTCACTCTAAGATTGATCCACATTTTACTGTAAGCAGAAACTTAGGACCCAAGATTATAA	2590					
TGCTGCAGAGGAI GCTGATCGATGTGTCTTCTCTCTGTTCTCTTTGC3GTGTGGATGGTGGCCTTTGG	2660					
CGTGGCCAGGCAAGGGATCCTTAGGCAGAA TGAGCAGCGCTGGAGGTGGATATTCCGTTGGGTGATCTAC	2730					
GAGCCCTACCTGGCCATGTTCCGCCAGGTGCCAGTGACGTGGATGGTACCACGTATGACTTGCCCACT	2800					
2810	2820	2830	2840	2850	2860	2870
GCACCTTCACTGGGAATGAGTCCAAGCCACTGTGTGTGGAGCTGGATGAGCACAACCTGCCCGGTTCCT	2870					
CGAGTGGATCACCATCCCCCTGGTGTGCATCTACATGTTATCCACCAACATCCTGCTGGTCAACCTGCTG	2940					
GTCGCCATGTTTGGCTACACGGTGGGCACCGTCCAGGAGAACAA TGACCAGGTCTGGAAGTCCAGAGGT	3010					
ACTTCTTGGTGCAGGAGTACTGCAGCCGCTCAATATCCCTTCCCTTCATCGTCTTGGCTTACTTCTA	3080					
CATGGTGGTGAAGAAGTGCTTCAAGTGTGTGCAAGGAGAAACATGGAGTCTTCTGTCTGTGTGTTT	3150					
3160	3170	3180	3190	3200	3210	3220
AAAAATGAAAGACAATGAGACTCTGGCATGGGAGGGTGTGATGAAGGAAAACCTACCTTGTCAAGATCAACA	3220					
CAAAAAGCCAAAGCAGACCTTCAGAGGAAA TGAGGCATCGATTTAGACAACCTGGATACAAAGCTTAATGATCT	3290					
CAAGGGTCTTCTGAAAGAGATTGCTTAATAAAATCAAAATAAACTGTATGAAACTCTAATGGAGAAAATC	3360					
TAATTATAGCAAGATCATATTAAGGAATGCTGATGAACATTTTGTATCGACTACTAAATGAGAGATTT	3430					
TCAGACCCCTGGGTACATGGTGGATGATTTTAAATCAGCTAGTGTGCTGAGACCTTGAGAATAAAGTGT	3500					
3510	3520	3530	3540	3550	3560	3570
GTGATTGGTTTCACTACTTGAAGACGGATATAAAGGAAAGATTTTCTTTATGTGTTCTCCAGAAATGGT	3570					
GGCTGTTTCTCTCTGTCTCTCAATGGCTGGGACTGGAGGTTGATAGTTTAAAGTGTGTCTTACCGCCTCC	3640					
TTTTTCTTTAATCTTATTTTGGATGAACACAATATAGGAGAACATCTATCTATGAATAAGAACCTGG	3710					
TCATGCTTACTCCTGTATTGTATTTTGTTCATTCGAATGGATTCTCTACTTTTCCCTTTTGTATT	3780					
ATGTGACTAATAGTTGGCATAATGTTAAAGTCTCTCAAAATAGGCCAGATTCTAAAACATGCTGCAGC	3850					
3860	3870	3880	3890	3900	3910	3920
AAGAGGACCCCGCTCTCTTTCAGGAAAAGTGTTCATTCTCAGGATGCTTCTTACCTGTGAGAGGAGGT	3920					
GACAAGGCAGTCTCTTGTCTCTTGGACTCAGCAGGCTCCTATTGAAGGAACACCCCAATTCCTAAATA	3990					
TGTGAAGAAGTCGCCAAAAATGCAACCTTGAAAGGCAGTACTGACTTTGTTCTTATTGGA TACTCCTCTTA	4060					
TTTATTATTTTTCATTAAAAAATAAGCTGGCTATTAGAAAAATTTAGACCATACAGAGATGTAGAAA	4130					
GAACATAAATTTTCCCCATTACCTTAAGGTAAACAGTGTAAACAATTTCTGGATGGTTTTTCAAGTCTAT	4200					
4210	4220	4230	4240	4250	4260	4270
TTTTTTCTATGATGTCTCAATCTCTCTTCAAAAATTTACAGAAATGTTATCATACTACATATATACTTT	4270					
TTATGTAAGCTTTTCACTTAGTATTTTATCAAAATATGTTTTATTATATTCATAGCCTTCTAAACATT	4340					
ATATCAATAAATTGCATAATAGGCAACCTCTAGCGATTACCAATAATTTTGTCTATTGAAGGCTATCTCCAG	4410					
TTGATCATTTGGATGAGCATCTTTGTGCATGAATCTATTGCTGTATTTGGGAAAAATTTTCCAAAGGTTAG	4480					
ATTCCAATAAATACTATTTATTATTAATAATTAATAATCGATTTATTAATAAACCAATTTATAAGGC	4550					

Fig. 12A(2)



4560	4570	4580	4590	4600	4610	4620
.....						
TTTTCATAAATGTATAGCAAAAGGAATTATTAACCTTGAGCATAAGATATGAGATACATGAACCTGAACT						4620
ATTAAAAATAAAATATTATATTTAACCCTAGTTTAAGAAGAAGTCAATATGCTTATTTAAATATTATGGAT						4690
GGTGGGCAGATCACTTGAGGTGAGGAGTTCGAGACCAGCCTGGCCAACATGGCAAAACCACATCTCTACT						4760
AAAAATAAAAAAATAGCTGGGTGTGGTGGTGCCTCTGTAATCCCAGCTACTCAGAAGGCTGAGGTAC						4830
AAGAATTGCTGGAACCTGGGAGGCGGAGGTTGCAGTGAACCAAGATTGCACCACTGCACTCCAGCCGGGG						4900

4910	4920	4930	4940	4950	4960	4970
.....						
TGACAGAGTGAGACTCCGACTGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATATTATGG						4970
ATGGTGAAGGGAATGGTATAGAATTGGAGAGATTATCTTACTGAACACCTGTAGTCCCAGCTTTCTCTGG						5040
AAGTGGTGGTATTTGAGCAGGATGTGCACAAGGCAATTGAAATGCCATAATTAGTTTCTCAGCTTTGAA						5110
TACACTATAAACTCAGTGGCTGAAGGAGGAAATTTAGAAGGAAGCTACTAAAAGATCTAATTTGAAAAA						5180
CTACAAAAGCATTAACTAAAAAGTTTATTTTCCTTTGTCTGGGCAGTAGTGAAAAATAACTACTCACAA						5250

5260	5270	5280	5290	5300	5310	5320
.....						
CATTCACTATGTTTGCAAGGAATTAACACAAATAAAAGATGCCTTTTTACTTAAACGCCAAGACAGAAAA						5320
CTTGCCCAATACTGAGAAGCAACTTGCATTAGAGAGGGAAGTGTAAATGTTTTCAACCCAGTTTCATCTG						5390
GTGGATGTTTTTGCAGGTACTCTGAGAATTTGCTTATGAAAAATCATTATTTTTAGTGTAGTTTACAA						5460
TAATGTATTGAACATACTTCTAATCAAAGGTGCTATGTCCTGTGTATGGTACTAAATGTGTCTGTGTGTA						5530
CTTTTGCACAACCTGAGAATCCTGCGGCTTGGTTTAAAGAGTGTGTTTCATGAAATAAATAATGGAGGAATT						5600

5610	5620	5630	5640	5650	5660	5670
.....						
GTCAA						5668

Fig. 12A(3)



10	20	30	40	50	60	70	
MRNRRNOTLOSTRILYSSASRSTOLSYSESOLVNF IQANFKKRECVFFTKDSKATENVCKCGYAQSQHME							70
GTQINQSEKWNYYKKHTKEFPTDAFGDIQFETLGKKGKYIRLSCOTDAEILYELLTQHWHLKTPNLVISVT							140
GGAKNFALKPRMRKIFSRLIYIAQSKGAWILTGGTHYGLTKYIGEVVRONTISRSSEENIVAIGIAAWGM							210
VSNRDTLIRNCDAEGYFLAQYLMDOFTROPLYLONNHTHLLLVDNGCHGHPTVEAKLRNQLEKHSERT							280
IQDSNYGGKIPVCFAGGGGKETLKAINTSINKKIPCVVVEGSGRIADVIAASLVEVEDAPTSSAVKEKLV							350
360	370	380	390	400	410	420	
RFLPRTVSR LSEEETESWIKWLKEILECSHLLTVIKMEZAGDEIVSNAISYALYKAFSTSEQOKONWNGQ							420
LKLLLEWNLQDLANDEIFTNORRWESADLQEVMTALIKDRPKFYRLFLENGLNLRKFLTHOVLTELF3N							490
HFSTLVYRNLGIAKNSYNDALLTFVWKL VANFRRGFRKEDRNGROEMOIELHGVSPITRHPLQALFIWA							560
LQNKKELSKVIWECTRGCTLAALGASKLLKTLAKYKNDINAAGESEELANEYETRAVELFTECYSSOEDL							630
AEQLLVYSCEAWGGSNCLELAVEATDQHFTAQPGVQNFLSKQWYGEISROTKNWKIILCLFIIPLVGCGF							700
710	720	730	740	750	760	770	
VSFRKKRPVQKHKLLWYYVAFFTSPFVVFVSWNVVFYIAFLLLFAYVLLMDFHSPHPPELVLYSLVFVLF							770
CDEVROQWYVNGVNYFTDLWNVMOTLGLFYFIAGIVFRHSSNKSSLYSGRYIFCLDYIFTLRLIHIFTV							840
SRNLGPKIIMLQRM LIDVFFFLFLFAYVMVAFGVARCGILRQNEQRWRIFRSVIYEPYLA MFGQVPSOV							910
DGTTYDFAHCTFTGNESKPLQVELDEHNLPRFPEWITIPLYCIYMLSTNILLVNL LVAMFGYTVGTVCEN							980
NDCVWKFQRYFLVQEYCSRLNIPFPFIVFAYFMVYKKCFKCCCKEKNMESSVCCFKNEDNETLAWEGVM							1050
1060	1070	1080	1090	1100	1110	1120	
KENYLVKINTKANOTSEEMRHRFRQLDTKLNLKGLKEIANKIK							1096

Fig. 12B



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Jiangchun Xu et al.  
Filed : June 13, 2000  
For : COMPOSITIONS AND METHODS FOR THERAPY AND  
DIAGNOSIS OF PROSTATE CANCER

Docket No. : 210121.42715C15

Date : June 13, 2000

Box Patent Application  
Assistant Commissioner for Patents  
Washington, D.C. 20231

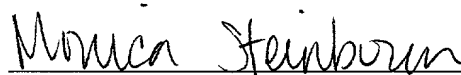
DECLARATION

Sir:

I, Monica Steinborn, in accordance with 37 C.F.R. § 1.821(f) do hereby declare that, to the best of my knowledge, the content of the paper entitled "Sequence Listing" and the computer readable copy contained within the floppy disk are the same.

I declare further that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Dated this 13<sup>th</sup> day of June, 2000.



Monica Steinborn  
Legal Assistant

701 Fifth Avenue, Suite 6300  
Seattle, WA 98104-7092  
(206) 622-4900  
FAX (206) 682-6031



## SEQUENCE LISTING

<110> Xu, Jiangchun  
 Dillon, Davin C.  
 Mitcham, Jennifer L.  
 Harlocker, Susan L.  
 Jiang, Yuqi  
 Reed, Steven G.  
 Kalos, Michael D.  
 Fanger, Gary R.  
 Retter, Marc W.  
 Stolk, John A.  
 Day, Craig H.  
 Vedvick, Thomas S.  
 Carter, Darrick  
 Li, Samuel  
 Wang, Aijun  
 Skeiky, Yasir A.W.  
 Helper, William  
 Henderson, Robert A.

<120> COMPOSITIONS AND METHODS FOR THERAPY AND  
 DIAGNOSIS OF PROSTATE CANCER

<130> 210121.42715C15

<140> US

<141> 2000-06-13

<160> 814

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 814

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(814)

<223> n = A,T,C or G

<400> 1

tttttttttt	tttttcacag	tataacagct	ctttatttct	gtgagttcta	ctaggaaatc	60
atcaaactctg	aggggttgtct	ggaggacttc	aatacacctc	cccccatagt	gaatcagctt	120
ccaggggggtc	cagtccctct	ccttacttca	tccccatccc	atgccaaagg	aagaccctcc	180
ctccttggct	cacagccttc	tctaggcttc	ccagtgcctc	caggacagag	tgggttatgt	240
tttcagctcc	atccttgctg	tgagtgtctg	gtgcgttgtg	cctccagctt	ctgctcagtg	300
cttcatggac	agtgtccagc	acatgtcact	ctccactctc	tcagtgtgga	tccactagtt	360
ctagagcggc	cgccaccgcg	gtggagctcc	agcttttgtt	cccttttagtg	agggttaatt	420



```

gcgcgcttgg cgtaatcatg gtcataactg tttcctgtgt gaaattgtta tccgctcaca      480
attccacaca acatacgagc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg      540
anctaactca cattaattgc gttgcgctca ctgnccgctt tccagtcngg aaaactgtcg      600
tgccagctgc attaatgaat cggccaacgc ncgggggaaaa gcggtttgcg ttttgggggc      660
tcttcgctt ctcgctcact nantcctgcg ctcggtcntt cggctgcggg gaacgggtatc      720
actcctcaaa gngggtatta cggttatccn naaatcnggg gatacccngg aaaaaanttt      780
aacaaaaggg cancaaaggg cngaaacgta aaaa                                     814

```

<210> 2

<211> 816

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(816)

<223> n = A,T,C or G

<400> 2

```

acagaaatgt tggatggtgg agcacctttc tatacgactt acaggacagc agatggggaa      60
ttcatggctg ttggagcaat agaacccagc ttctacgagc tgctgatcaa aggacttgga      120
ctaaagtctg atgaacttcc caatcagatg agcatggatg attggccaga aatgaagaag      180
aagtttgtag atgtatttgc aaagaagacg aaggcagagt ggtgtcaaat ctttgacggc      240
acagatgcct gtgtgactcc ggttctgact tttgaggagg ttgttcatca tgatcacaac      300
aaggaacggg gctcgtttat caccagttag gagcaggacg tgagcccccg ccctgcacct      360
ctgctgttaa acacccagc catcccttct ttcaaaaggg atccactagt tctagaagcg      420
gccgccaccg cgggtggagct ccagcttttg ttccctttag tgagggttaa ttgcgcgctt      480
ggcgtaatca tggatcatagc tgtttcctgt gtgaaattgt tatccgctca caattcccc      540
aacatacgag ccggaacata aagtgttaag cctgggggtgc ctaatgantg agctaactcn      600
cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaaactgtcg tgccactgcn      660
ttantgaatc ngccaccccc cgggaaaagg cggttgcntt ttgggcctct tccgctttcc      720
tcgctcattg atcctngcnc ccggtcttcg gctgcggnga acggttcact cctcaaaggc      780
ggtntnccgg ttatccccaacn cngggggata cccnga                                     816

```

<210> 3

<211> 773

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(773)

<223> n = A,T,C or G

<400> 3

```

cttttgaaaag aagggatggc tgggggtgttt aacagcagag gtgcagggcg ggggctcacg      60
tcttgctcct cactggtgat aaacgagccc cgttccttgt tgtgatcatg atgaacaacc      120
tctcaaaaag tcagaaccgg agtcacacag gcatctgtgc cgtcaaagat ttgacaccac      180
tctgccttcg tcttctttgc aaatacatct gcaaacttct tcttcatttc tggccaatca      240
tccatgctca tctgattggg aagttcatca gactttagtc canntccttt gatcagcagc      300
tcgtagaact ggggttctat tgctccaaca gccatgaatt ccccatctgc tgtcctgtaa      360

```



```

gtcgtataga aaggtgctcc accatccaac atgtttctgtc ctcgaggggg ggccccggtac 420
ccaattcgcc ctatantgag tcgtattacg cgcgctcact ggccgctcgtt ttacaacgtc 480
gtgactggga aaaccctggg cgttaccaac ttaatcgcct tgcagcacat ccccttttcg 540
ccagctgggc gtaatanca aaaggccgc accgatcgcc cttccaacag ttgcgcacct 600
gaatgggnaa atgggacccc cctgttaccg cgcattnaac cccgcnggg tttngttgtt 660
acccccacnt nnaccgctta cactttgcc a gcgccttanc gccgctccc tttcnccttt 720
cttccttcc tttcncncn ctttcccccg ggggttcccc cntcaaacc cna 773

```

```

<210> 4
<211> 828
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(828)
<223> n = A,T,C or G

```

```

<400> 4
cctcctgagt cctactgacc tgtgctttct ggtgtggagt ccagggctgc taggaaaagg 60
aatgggcaga cacagggtgta tgccaatgtt tctgaaatgg gtataatttc gtcctctcct 120
tcggaacact ggctgtctct gaagacttct cgctcagttt cagtgaggac acacacaaaag 180
acgtgggtga ccatgttggt tgtgggggtgc agagatggga ggggtggggc ccaccctgga 240
agagtggaca gtgacacaag gtggacactc tctacagatc actgaggata agctggagcc 300
acaatgcatg aggcacacac acagcaagga tgacnctgta aacatagccc acgctgtcct 360
gngggcactg ggaagcctan atnaggccgt gagcanaaag aaggggagga tccactagtt 420
ctanagcggc cgccaccgcg gtgganctcc ancttttggt cccttttagtg agggttaatt 480
gcgcgcttgg cntaatcatg gtcatanctn tttcctgtgt gaaattgtta tccgctcaca 540
attccacaca acatacganc cggaacata aantgtaaac ctgggggtgcc taatgantga 600
ctaaactaca ttaattgctg tgcgctcact gcccgctttc caatcnggaa acctgtcttg 660
ccncttgcat tnatgaatcn gccaaccccc ggggaaaagc gtttgcgttt tgggcgctct 720
tccgcttcct cnetcantta ntceetncnc tcggtcatte cggctgcngc aaaccggttc 780
accncctcca aaggggggtat tccggtttcc ccnaatccgg gganancc 828

```

```

<210> 5
<211> 834
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(834)
<223> n = A,T,C or G

```

```

<400> 5
tttttttttt tttttactga tagatggaat ttattaagct tttcacatgt gatagcacat 60
agttttaatt gcatccaaag tactaacaaa aactctagca atcaagaatg gcagcatggt 120
attttataac aatcaacacc tgtggcctttt aaaatttggt tttcataaga taatttatac 180
tgaagtaaat ctagccatgc ttttaaaaaa tgcttttaggt cactccaagc ttggcagtta 240
acatttgcca taaacaataa taaaacaatc acaattttaat aaataacaaa tacaacattg 300
taggccataa tcatatacag tataaggaaa aggtggtagt gttgagtaag cagttattag 360

```



```

aatagaatac cttggcctct atgcaaatat gtctagacac tttgattcac tcagccctga 420
cattcagttt tcaaagtagg agacagggttc tacagtatca ttttacagtt tccaacacat 480
tgaaaacaag tagaaaatga tgagttgatt tttattaatg cattacatcc tcaagagtta 540
tcaccaaccc ctcagttata aaaaattttc aagttatatt agtcatataa cttgggtgtgc 600
ttatttttaa ttagtgctaa atggattaag tgaagacaac aatgggtccc taatgtgatt 660
gatattgggtc atttttacca gcttctaaat ctnaactttc aggcttttga actggaacat 720
tgnatnacag tgttccanag ttncaaccta ctggaacatt acagtgtgct tgattcaaaa 780
tgttattttg ttaaaaatta aattttaacc tgggtggaaa ataatttgaa atna 834

```

```

<210> 6
<211> 818
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(818)
<223> n = A,T,C or G

```

```

<400> 6
tttttttttt tttttttttt aagaccctca tcaatagatg gagacataca gaaatagtca 60
aaccacatct acaaaatgcc agtatcaggc ggcggcttcg aagccaaagt gatgtttgga 120
tgtaaagtga aatattagtt ggcggatgaa gcagatagtg aggaaagttg agccaataat 180
gacgtgaagt ccgtggaagc ctgtggctac aaaaaatgtt gagccgtaga tgccgtcggg 240
aatggtgaag ggagactcga agtactctga ggctttagtg agggtaaaat agagaccag 300
taaaattgta ataagcagtg cttgaattat ttggtttcgg ttgttttcta ttagactatg 360
gtgagctcag gtgattgata ctctgatgc gagtaatacg gatgtgttta ggagtgggac 420
ttctagggga ttttagcggg tgatgcctgt tgggggccag tgccctccta gttggggggg 480
aggggctagg ctggagtggg aaaaggctca gaaaaatcct gcgaagaaaa aaacttctga 540
ggtaataaat aggattatcc cgtatcgaag gccttttttg acagggtggg tggtgtggcc 600
ttggtatgtg ctttctcgtg ttacatcgcg ccatcattgg tatatgggta gtgtgtggg 660
ttantangg ctantatgaa gaacttttg antggaatta aatcaatngc ttggccggaa 720
gtcattanga nggctnaaaa ggccctgtta ngggtctggg ctnggtttta cccnaccat 780
ggaatncncc ccccggaacna ntgnatccct attcttaa 818

```

```

<210> 7
<211> 817
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(817)
<223> n = A,T,C or G

```

```

<400> 7
tttttttttt tttttttttt tggctctaga gggggtagag ggggtgctat agggtaaata 60
cgggccctat ttcaaagatt tttaggggaa ttaattctag gacgatgggt atgaaactgt 120
ggtttgctcc acagatttca gagcattgac cgtagtatac ccccggtcgt gtagcgggta 180
aagtggtttg gtttagacgt ccgggaattg catctgtttt taagcctaata gtggggacag 240
ctcatgagtg caagacgtct tgtgatgtaa ttattatacn aatgggggct tcaatcggga 300

```



```

gtactactcg attgtcaacg tcaaggagtc gcaggtcgcc tggttctagg aataatgggg 360
gaagtatgta ggaattgaag attaatccgc cgtagtcggt gttctcctag gttcaatacc 420
attggtggcc aattgatttg atggttaagg gagggatcgt tgaactcgtc tggtatgtaa 480
aggatncctt ngggatggga aggcnatnaa ggactangga tnaatggcgg gcangatatt 540
tcaaacngtc tctanttcct gaaacgtctg aaatgttaat aanaattaan tttngttatt 600
gaatnttnng gaaaagggct tacaggacta gaaaccaa atnntaangg 660
cnttatcntn aaaggtnata accnctocta tnatccacc caatngnatt cccacnchn 720
acnattggat nccccanttc canaaanggc cccccccgg tgnannccnc cttttgttcc 780
cttnantgan ggttattcnc cctngcntt atcance 817

```

<210> 8

<211> 799

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(799)

<223> n = A,T,C or G

<400> 8

```

catttcgggg tttactttct aaggaaagcc gagcggaagc tgctaacgtg ggaatcggtg 60
cataaggaga actttctgct ggcacgcgct agggacaagc gggagagcga ctccgagcgt 120
ctgaagcgca cgtcccagaa ggtggacttg gcaactgaaac agctgggaca catccgcgag 180
tacgaacagc gcctgaaagt gctggagcgg gaggtccagc agtgtagccg cgtcctgggg 240
tggttgggcc angcctganc cgctctgctt tgctgcccc angtgggccg ccacccctg 300
acctgcctgg gtccaaacac tgagccctgc tggcggactt caagganaac cccacangg 360
ggattttgct cctanantaa ggctcatctg ggctcggcc cccacactg gttggccttg 420
tctttgangt gagccccatg tccatctggg ccaactgtcng gaccaccttt ngggagtgtt 480
ctccttacia ccacannatg cccggctcct cccggaacc antccancc tnggaaggat 540
caagnccctg atccactnnt nctanaaccg gccnccnccg cngtggaacc cnccttntgt 600
tccttttcnt tnaggggttaa tnnccgcttg gccttnccan ngctcncnc ntttccnnt 660
gttnaaattg ttangcnccc nccnntccn cnnnncnan cccgaccnn annttnnann 720
ncctgggggt nccnncgat tgaccnnc nccctntant tgcnttnggg nncnntgccc 780
ctttccctct nggganncg 799

```

<210> 9

<211> 801

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(801)

<223> n = A,T,C or G

<400> 9

```

acgccttgat cctcccaggc tgggactggt tctgggagga gccgggcatg ctgtggtttg 60
taangatgac actcccaaag gtggctcctga cagtggccca gatggacatg gggctcacct 120
caaggacaag gccaccaggt gcggggggccg aagcccacat gatccttact ctatgagcaa 180
aatccctgtt gggggcttct ccttgaagtc cgccancagg gctcagctt tggacccang 240

```







```

tgtgggctga ggggacctgg ttcttgtgtg ttgcccctca ggactcttcc cctacaaata      240
actttcatat gttcaaatcc catggaggag tgtttcatcc tagaaactcc catgcaagag      300
ctacattaaa cgaagctgca ggttaagggg cttanagatg ggaaaccagg tgactgagtt      360
tattcagctc ccaaaaaccc ttctctaggt gtgtctcaac taggaggcta gctgttaacc      420
ctgagcctgg gtaatccacc tgcagagtc cgcattcca gtgcatggaa cccttctggc      480
ctccctgtat aagtccagac tgaaccccc ttggaaggnc tccagtcagg cagccctana      540
aactggggaa aaaagaaaag gacgccccan ccccagctg tgcantacg cacctcaaca      600
gcacaggggtg gcagcaaaaa aaccacttta ctttggcaca aacaaaaact ngggggggca      660
accccggcac cccnangggg gttaacagga ancngggnaa cntggaaccc aattnaggca      720
ggcccnccac ccnaatntt gctgggaaat ttttctccc ctaaattntt tc              772

```

```

<210> 12
<211> 751
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(751)
<223> n = A,T,C or G

```

```

<400> 12
gccccaatc cagctgccac accaccacg gtgactgcat tagttcggat gtcatacaaa      60
agctgattga agcaaccctc tacttttttg tcgtgagcct tttgcttggg gcaggtttca      120
ttggctgtgt tgggtgacgtt gtcattgcaa cagaatgggg gaaaggcact gttctctttg      180
aagtanggtg agtcctcaaa atccgtatag ttggtgaagc cacagcactt gagccctttc      240
atggtggtgt tccacacttg agtgaagtct tcctgggaac cataatcttt cttgatggca      300
ggcactacca gcaacgtcag ggaagtgctc agccattgtg gtgtacacca aggcgaccac      360
agcagctgcn acctcagcaa tgaagatgan gaggangatg aagaagaacg tcncgagggc      420
acacttgctc tcagtcttan caccatanca gcccntgaaa accaananca aagaccacna      480
cnccggctgc gatgaagaaa tnaccccnng ttgacaaact tgcatggcac tggganccac      540
agtggcccnna aaaatcttca aaaaggatgc cccatcnatt gaccccccaa atgccactg      600
ccaacagggg ctgccccacn cncnnaacga tgancnatt gnacaagatc tncntggtct      660
tnatnaacnt gaacctgcn tngtggctcc tgttcaggnc cnnggcctga cttctnaann      720
aangaactcn gaagncccca cngganann c g              751

```

```

<210> 13
<211> 729
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(729)
<223> n = A,T,C or G

```

```

<400> 13
gagccaggcg tccctctgcc tgcccactca gtggcaacac ccgggagctg ttttgcctt      60
tgtggancct cagcagtncc ctctttcaga actcantgcc aaganccctg aacaggagcc      120
accatgcagt gcttcagctt cattaagacc atgatgatcc tcttcaattt gtcacatctt      180
ctgtgtgggtg cagccctgtt ggcagtgggc atctgggtgt caatcgatgg ggcacatctt      240

```



```
<210> 14
<211> 816
<212> DNA
<213> Homo sapien
```

<400> 14

```
<210> 15
<211> 783
<212> DNA
<213> Homo sapien
```

<400> 15

ccaaggcctg	ggcaggcata	nacttgaagg	tacaacccca	ggaacccctg	gtgctgaagg	60
atgtggaaaa	cacagattgg	cgctactgc	gggtgacac	ggatgtcagg	gtagagagga	120
aagacccaaa	ccaggtggaa	ctgtggggac	tcaaggaang	cacctacctg	ttccagctga	180
cagtgactag	ctcagaccac	ccagaggaca	cggccaacgt	cacagtcact	gtgctgtcca	240



```
<210> 16
<211> 801
<212> DNA
<213> Homo sapien
```

<400> 16

```
<210> 17
<211> 740
<212> DNA
<213> Homo sapien
```

<400> 17

```
gtgagagcca ggcgtccctc tgcctgccca ctcagtggca acacccgga gctgttttgt      60
cctttgtgga gcctcagcag ttccctcttt cagaactcac tgccaagagc cctgaacagg      120
agccaccatg cagtgotcca gcttcattaa gaccatgatg atcctcttca atttgctcat      180
```



```

ctttctgtgt ggtgcagccc tgttggcagt gggcatctgg gtgtcaatcg atggggcatc 240
ctttctgaag atcttcgggc cactgtcgtc cagtgccatg cagtttgtca acgtgggcta 300
cttcctcatc gcagccggcg ttgtggtcct tgcctcttgg ttccctgggct gctatgggtgc 360
taagacggag agcaagtgtg ccctcgtgac gttcttcttc atcctcctcc tcatcttcat 420
tgctgaagtt gcagctgctg tggcgcctt ggtgtacacc acaatggctg aaccattcct 480
gacgttgctg gtantgcctg ccatcaanaa agattatggg ttcccaggaa aaattcactc 540
aantntggaa caccnccatg aaaagggctc caatttctgn tggcttcccc aactataccg 600
gaattttgaa agantcnccc tacttccaaa aaaaaanant tgccttttnc cccnttctgt 660
tgcaatgaaa acntcccaan acngccaatn aaaacctgcc cnnncaaaaa ggntcncaaa 720
caaaaaaant nnaagggttn 740

```

```

<210> 18
<211> 802
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(802)
<223> n = A,T,C or G

```

```

<400> 18
ccgctggttg cgctggtcca gngnagccac gaagcacgtc agcatacaca gcctcaatca 60
caaggtcttc cagctgccgc acattacgca gggcaagagc ctccagcaac actgcatatg 120
ggatacactt tacttttagca gccagggtga caactgagag gtgtcgaagc ttattcttct 180
gagcctctgt tagtggagga agattccggg cttcagctaa gtagtcagcg tatgtcccat 240
aagcaaacac tgtgagcagc cggaaggtag aggcaaagtc actctcagcc agctctctaa 300
cattgggcat gtccagcagt tctccaaaca cgtagacacc agnggcctcc agcacctgat 360
ggatgagtgt ggccagcgct gcccccttgg ccgacttggc taggagcaga aattgctcct 420
ggttctgccc tgtcaccttc acttcgcgac tcatcactgc actgagtgtg ggggacttgg 480
gctcaggatg tccagagacg tggttccgcc ccctcnctta atgacaccgn ccanncaacc 540
gtcggctccc gccgantgng ttcgctcgtc ctgggtcagg gtctgctggc cinctacttgc 600
aancttcgtc nggccccatg aattcaccnc accggaactn gtangatcca ctnnttctat 660
aaccggncgc caccgcnntt ggaactccac tcttnttnc tttacttgag ggtaaggtc 720
acccttnncc ttaccttggt ccaaaccntn ccntgtgtcg anatngtnaa tcnggncna 780
tnccancnc atangaagcc ng 802

```

```

<210> 19
<211> 731
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(731)
<223> n = A,T,C or G

```

```

<400> 19
cnaagcttcc aggtnacggg ccgcnaance tgaccnagg tancanaang cagncngcgg 60
gagcccaccg tcacngggng gngtctttat nggagggggc ggagccacat cnetggacnt 120
cntgacccca actcccncc ncnantgca gtgatgagtg cagaactgaa ggtnacgtgg 180

```



```

caggaaccaa gancaaannc tgetccnntc caagtcggcn nagggggcgg ggctggccac 240
gncatccnt cnagtgetgn aaagcccenn cctgtctact tgtttgaga acngcnnga 300
catgccagn gttanataac nggngagag tnannttgcc tctccctcc ggctgcgan 360
cngtntgct tagnggacat aacctgacta cttaactgaa ccnngaate tncnccccct 420
ccactaagct cagaacaaaa aacttcgaca ccactcantt gtcacctgnc tgctcaagta 480
aagtgtaccc catncccaat gtntgctnga ngctctgncc tgcnttangt tcggtcctgg 540
gaagacctat caattnaagc tatgtttctg actgcctctt gctccctgna acaancnacc 600
cnncnntcca aggggggggnc ggcccccaat ccccccaacc ntnaattnan tttanccccn 660
ccccnggcc cggcctttta cnanctcnn nnacngggna aaaccnnngc tttncccaac 720
nnaatccncc t 731

```

```

<210> 20
<211> 754
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (754)
<223> n = A,T,C or G

```

```

<400> 20
tttttttttt tttttttttt taaaaacccc ctccattnaa tgnaaacttc cgaaattgtc 60
caacccctc ntccaaatnn cnttttcgg gnggggggttc caaacccaan ttanntttgg 120
annttaaatt aaatnttnnt tggnggnna anccnaatgt nangaaagt naaccanta 180
tnancttnaa tncctgaaa cngtngntt ccaaaaatnt ttaaccetta antccctccg 240
aaatngttna nggaaaaccc aanttctcnt aaggttggtt gaaggntnaa tnaaaanccc 300
nnccaattgt tttngccac gcctgaatta attggnntcc gntgttttcc nttaaaanaa 360
ggnnancccc gggtantnaa tcccccnnc cccaattata ccganttttt ttngaattgg 420
gancccnccg gaattaacgg ggnnnntccc tnttgggggg cnggnncccc cccntcggg 480
ggttngggnc aggnccnaat tgtttaaggg tccgaaaaat ccctccnaga aaaaaanctc 540
ccaggntgag nntnggggtt ncccccccc canggccct ctcgnanagt tggggtttgg 600
ggggcctggg attttntttc ccctnttncc tcccccccc ccnggganag aggttngngt 660
tttgntcnnc ggcccccncc aaganccttn ccganttnan ttaaaccnt gcctnggcga 720
agtcctntgn agggntaaan ggccccctnn cggg 754

```

```

<210> 21
<211> 755
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (755)
<223> n = A,T,C or G

```

```

<400> 21
atcancccat gaccccnac nngggaccnc tcanceggnc nnncnaccnc cggccnatca 60
nngtnagnnc actncnnttn natcacnccc cncnactac gcccnananc cnacgcncta 120
nncanatncc actganngcg cgangtngan ngagaaanct nataccanag ncaccanacn 180
ccagctgtcc nanaangcct nnnatacnng nnnatccaat ntgnancctc cnaagtattn 240

```







```

nggcgaatcg taatnaggcg tgcgcgcga atntgtcncc gtttatntn ccagcntcnc 240
ctnccnacc tacntcttcn nagctgtcnn acccctngtn cgnaccccc naggtcggga 300
tcgggtttnn nntgaccgng cnnccctcc cccctccat nacganccnc ccgcaccacc 360
nanngcncgc nccccgnnct cttocecncc ctgtcctntn cccctgtngc ctggcncngn 420
accgcattga ccctcgccnn ctncnngaaa ncnanacgt ccgggttggn annancgctg 480
tgggnnngcg tctgncgcgc gttccttcn ncncttcca ccatcttcnt tacngggctc 540
ccncgcctc tcnnncacnc cctgggacgc tntcctntgc ccccttnac tccccctt 600
cgcgtgncc cgnccccacc ntcatttnca nacgntcttc acaannncct ggntnnctcc 660
cnancngncn gtcancnag ggaagggngg ggnncnntg nttgacgttg ngngangtc 720
cgaanantcc tcncntcan cctaccct cgggcgnnct ctngttnc aacttancaa 780
ntctcccccg ngngcncntc tcagcctcnc cccccnct ctctgcantg tntctgctc 840
tnaccnntac gantnttcgn cncctcttt cc 872

```

```

<210> 24
<211> 815
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(815)
<223> n = A,T,C or G

```

```

<400> 24
gcatgcaagc ttgagtattc tatagngtca cctaaatanc ttggcntaat catggtcnta 60
nctgncttcc tgtgtcaaata gtatacnaan tanatatgaa tctnatntga caaganngtg 120
tcntncatta gtaacaantg tnntgtccat cctgtcngan canattccca tnnattncgn 180
cgcattcncn gcncantatn taatngggaa ntcnnntnnn ncaccnncat ctatcntncc 240
gcnccttgac tggagagat ggatnanttc tnntntgacc nacatgttca tcttggattn 300
aanaccccc cgcngnccac cggttngnng cnagccnntc ccaagacctc ctgtggaggt 360
aacctgcgtc aganncatca aacntgggaa acccgcnnc angtnnaagt ngnnncanan 420
gatcccgctc aggnntnacc atcccttcnc agcgccccct ttngtgcctt anagnnagc 480
gtgtccnanc cnetcaacat ganacgcgcc agnccanccg caattnggca caatgtcngc 540
gaacccccct gggggantna tncaaanccc caggattgtc cncncangaa atccncanc 600
cccncctac ccncttttg gacngtgacc aantcccgga gtncagtc gcngnctc 660
ccccaccggt nncntgggg ggggtgaanct cngnntcanc cngncgaggn ntcgnaagga 720
accggnccn ggncgaanng ancnntcnga agngcncnt cgtataacc cccctcncca 780
nccnancngt agntcccccc cngggtncgg aangg 815

```

```

<210> 25
<211> 775
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(775)
<223> n = A,T,C or G

```

```

<400> 25
ccgagatgtc togetccgtg gccttagctg tgctcgcgt actctctctt tctggcctgg 60

```



```

aggctatcca gcgtactcca aagattcagg tttactcacg tcatccagca gagaatggaa 120
agtcaaattt cctgaattgc tatgtgtctg ggtttcatcc atccgacatt gaanttgact 180
tactgaagaa tgganagaga attgaaaaag tggagcattc agacttgtct ttcagcaagg 240
actggtcttt ctatctcntg tactacactg aattcacccc cactgaaaaa gatgagtatg 300
cctgccgtgt gaaccatgtg actttgtcac agcccaagat agttaagtgg gatcgagaca 360
tgtaagcagn cnnecatggaa gtttgaagat gccgcatttg gattggatga attccaaatt 420
ctgcttgctt gcnttttaat antgatatgc ntatacacc taccctttat gnccccaaatt 480
tgtaggggtt acatnantgt tcnctnngga catgatcttc ctttataant ccncnttcg 540
aattgcccgt cncncngttn ngaatgtttc cnaaccacg gttggctccc ccaggtcncc 600
tcttacggaa gggcctgggc cnccttncaa ggttggggga accnaaaatt tcnctntngc 660
ccnccncca cnntcttgng nncncanttt ggaacccttc cnattcccct tggcctcnna 720
nccttnncta anaaaacttn aaancgtngc naaanntttn acttcccccc ttacc 775

```

```

<210> 26
<211> 820
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(820)
<223> n = A,T,C or G

```

```

<400> 26
anattantac agtgtaatct tttcccagag gtgtgtanag ggaacggggc ctagaggcat 60
cccanagata ncttatanca acagtgcctt gaccaagagc tgctgggcac atttcctgca 120
gaaaagggtg cggtccccc cactcctcct ctcccatagc catcccagag gggtgagtag 180
ccatcangcc ttcggtggga gggagtcang gaaacaacan accacagagc anacagacca 240
ntgatgacca tgggcgggag cgagcctctt ccctgnaccg gggtggcana nganagccta 300
nctgaggggt cacactataa acgttaacga ccnagatnan cacctgcttc aagtgcaccc 360
ttcctacctg acnaccagng accnnnaact gcngcctggg gacagcncct ggancagcta 420
acnnagcact cacctgcccc cccatggccg tncgntccc tggctcctgnc aaggggaagct 480
ccctgttgga attncgggga naccaaggga nccccctcct ccancctgtga aggaaaaann 540
gatggaattt tnccttccg gccnntcccc tcttccttta cagccccct nntactentc 600
tcctctntt ntcctgnenc acttttnacc ccnnnatttc ccttnattga tcggannctn 660
ganattccac tnngcctnc cntcnatcng naanacnaaa nactntctna ccnggggat 720
gggnncctcg ntcatectct ctttttctct accnccnntt ctttgectct ccttngatca
780tccaaccntc gntggcentn ccccccnnn tcctttncce 820

```

```

<210> 27
<211> 818
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(818)
<223> n = A,T,C or G

```

```

<400> 27
tctgggtgat ggcctcttcc tctcagga cctctgactg ctctgggcca aagaatctct 60

```







```

cgctcanacc tcacancctc ccnacnangc ctataangaa nannaataga nctgtncnnt 120
atntntacnc tcatanncct cnnnacccac tccctcttaa cccntactgt gcctatngcn 180
tnnctantct ntgcgcgctn cnanccaccn gtggggccnac cncnngnatt ctcnatctcc 240
tcnccatntn gcctananta ngtncatacc ctataacctac nccaatgcta nnnctaanch 300
tccatnantt annntaacta ccactgacnt ngactttcnc atnanctcct aatttgaatc 360
tactctgact cccacngcct annnattagc ancntcccc nacnatntct caaccaaadc 420
ntcaacaacc tatctanctg ttcnccaacc nttncctccg atccccnnac aacccccctc 480
ccaaataccc nccacctgac nccaaaccn caccatcccg gcaagccnan ggncatttan 540
ccactggaat cacnatngga naaaaaaac ccnaactctc tancncnnat ctccctaana 600
aatnctcctn naatttactn ncantnccat caanccacn tgaaacnaa cccctgtttt 660
tanatccctt ctttcgaaaa ccnacccttt annncccaac ctttngggcc ccccnctnc 720
ccnaatgaag gncncccaat cnangaaacg nccntgaaaa ancnaaggcna anannntccg 780
canatcctat cccttanttn ggggnccctt nccngggcc cc 822

```

```

<210> 30
<211> 787
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(787)
<223> n = A,T,C or G

```

```

<400> 30
cggcgcgctg ctctggcaca tgcctcctga atggcatcaa aagtgatgga ctgcccattg 60
ctagagaaga ccttctctcc tactgtcatt atggagccct gcagactgag ggctccccctt 120
gtctgcagga tttgatgtct gaagtogtgg agtgtggctt ggagctcctc atctacatna 180
gctggaagcc ctggagggcc tctctogcca gcctccccct tctctccacg ctctccangg 240
acaccagggg ctccaggcag cccattatct ccagnangac atgggtgtttc tccacgcgga 300
cccatggggc ctgnaaggcc aggggtctct ttgacaccat ctctcccgct ctgcctggca 360
ggcgtggga tccactantt ctanaacggg cgccaccncg gtgggagctc cagcttttgt 420
tccnttaaat gaaggttaat tgcncgcttg gcgtaatcat nggtcanaac tntttcctgt 480
gtgaaattgt ttntccccct ncnatccnc ncnacatacn aacccggaan cataaagtgt 540
taaagcctgg gggtngcctn nngaataaac tnaactcaat taattgcgtt ggctcatggc 600
ccgctttccn ttcnngaaaa ctgtcntccc ctgcnttnnt gaatcggcca ccccccnggg 660
aaaagcgggt tgcnttttng ggggntcctt ccncttcccc cctcnctaan cccnncgctt 720
cggtcgttnc nggtngcggg gaangggnat nnnctccnc naagggggng agnnngntat 780
ccccaaa 787

```

```

<210> 31
<211> 799
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(799)
<223> n = A,T,C or G

```

```

<400> 31

```



```

tttttttttt tttttttggc gatgctactg ttttaattgca ggaggtgggg gtgtgtgtac      60
catgtaccag ggctattaga agcaagaagg aaggagggag ggcagagcgc cctgctgagc      120
aacaaaggac toctgcagcc ttctctgtct gtctcttggc gcaggcacat ggggaggcct      180
cccgagggtt gggggccacc agtccagggg tgggagcact acanggggtg ggagtgggtg      240
gtggctggtt cnaatggcct gncacanatc cctacgattc ttgacacctg gatttcacca      300
ggggaccttc tgttctccca nggnaacttc nttnatctcn aaagaacaca actgtttctt      360
cngcanttct ggctgttcat ggaaagcaca ggtgtccnat ttnggctggg acttgggtaca      420
tatggttccg gcccacctct cccntcnaaa aagtaattca ccccccccn cctctntttg      480
cctgggccct taantacca caccggaact canttanta ttcattctng gntgggcttg      540
ntnatcnccn cctgaangcg ccaagttgaa aggccacgcc gtncnctc cccatagnan      600
nttttnnctn canctaagtc cccccnggc aacnatccaa tcccccccn tgggggcccc      660
agcccanggc ccccgncctg ggnnnccngn cncgnantcc ccaggntctc ccantcngnc      720
ccnnngcncc cccgcacgca gaacanaagg ntngagccnc cgcannnnnn nggtnncnac      780
ctgccccccc ccnnccgngg                                     799

```

```

<210> 32
<211> 789
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (789)
<223> n = A,T,C or G

```

```

<400> 32
tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt      60
ttttnccnag ggcagggtta ttgacaacct cncgggacac aancaggctg gggacaggac      120
ggcaacaggc tccggcgggc ggcggcgggc ccctacctgc ggtaccaaata ntgcagcctc      180
cgctcccgtt tgatnttccg ctgcagctgc aggatgccnt aaaacagggc ctcgcccntn      240
ggtgggcacc ctgggatttn aatttccacg ggcacaatgc ggtcgcancc cctcaccacc      300
nattaggaat agtggtnnta cccnccnccg ttggcncact ccccntggaa accactnttc      360
gcggctccgg catctggtct taaaccttgc aaacnctggg gccctctttt tggttantnt      420
nccngccaca atcatnactc agactggcnc gggctggccc caaaaaancn ccccaaaacc      480
ggnccatgtc ttnnccgggt tgctgcnatn tncatcacct cccgggcnca ncaggncaac      540
ccaaaagtgc ttngggcccn caaaaaanct cgggggggnc ccagtttcaa caaagtcatc      600
ccccttggcc cccaaatcct cccccgntt nctgggtttg ggaaccacag cctctnnctt      660
tggnnggcaa gntggntccc ccttcgggccc cccgggtggc ccnnctctaa ngaaaaacnc      720
ntcctnnnca ccatcccccc nngnnacgnc tancaangna tccctttttt tanaaacggg      780
ccccccnccg                                     789

```

```

<210> 33
<211> 793
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1) ... (793)
<223> n = A,T,C or G

```



&lt;400&gt; 33

```

gacagaacat gttggatggt ggagcacctt tctatacgac ttacaggaca gcagatgggg      60
aattcatggc tgttggagca atanaacccc agttctacga gctgctgac aaaggacttg      120
gactaaagtc tgatgaactt cccaatcaga tgagcatgga tgattggcca gaaatgaana      180
agaagtttgc agatgtattht gcaaagaaga cgaaggcaga gtggtgtcaa atctttgacg      240
gcacagatgc ctgtgtgact cgggttctga cttttgagga ggttggtcat catgatcaca      300
acaangaacg gggctcgttt atcaccantg aggagcagga cgtgagcccc cgccctgcac      360
ctctgctgtht aaacacccca gccatccctt ctttcaaaaag ggatccacta cttctagagc      420
ggncgccacc gcggtggagc tccagctttt gttcccttta gtgagggtta attgcgcgct      480
tggcgtaatc atggtcatan ctgtttcctg tgtgaaattg ttatccgctc acaattccac      540
acaacatacg anccggaagc atnaaatttt aaagcctggg ggtngcctaa tgantgaact      600
nactcacatt aattggcttt gcgctcactg cccgctttcc agtccggaaa acctgtcctt      660
gccagctgcc nttaatgaat cnggccaccc cccggggaaa aggcngtttg cttnttgggg      720
cgcnccttccc gctttctcgc ttctgaant ccttcccccc ggtctttcgg cttgcggcna      780
acggtatcna cct

```

&lt;210&gt; 34

&lt;211&gt; 756

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(756)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 34

```

gccgcgaccg gcatgtacga gcaactcaag ggcgagtgga accgtaaaaag ccccaatctt      60
ancaagtgcg ggggaanagct gggctgactc aagctagttc ttctggagct caacttcttg      120
ccaaccacag ggaccaagct gaccaaacag cagctaattc tggcccgtga catactggag      180
atcggggccc aatggagcat cctacgcaan gacatcccct ccttcgagcg ctacatggcc      240
cagctcaaat gctactactt tgattacaan gagcagctcc ccgagtcagc ctatatgcac      300
cagctcttgg gcctcaacct cctcttctctg ctgtcccaga accgggtggc tgantnccac      360
acgganttgg ancggctgcc tgccccanga catacanacc aatgtctaca tcnaccacca      420
gtgtccttga gcaatactga tgganggcag ctaccncaaa gtnttctctg ccnagggtta      480
catccccgcg cgagagctac accttcttca ttgacatcct gctcgacact atcagggatg      540
aaaatcgcn ggttgctcca gaaaggctnc aanaanatcc ttttcnctga aggcccccg      600
atncnctagt nctagaatcg gcccgccatc gcggtgganc ctccaacctt tcgttnccct      660
ttactgaggg ttnattgccg cccttggcgt tatcatggtc acncngttn cctgtgttga      720
aattnttaac cccccacaat tccacgcna cattng

```

&lt;210&gt; 35

&lt;211&gt; 834

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(834)

&lt;223&gt; n = A,T,C or G



ggggatctct	anatanacct	gnatgcatgg	ttgtcggtgt	ggtcgtgtc	gatgaanatg	60
aacaggatct	tgcccttgaa	gctctcggct	gctgtnttta	agttgctcag	tctgccgtca	120
tagtcagaca	cnctcttggg	caaaaaacan	caggatntga	gtcttgattt	cacctccaat	180
aatcttcnng	gctgtctgct	cgggtgaactc	gatgacnang	ggcagctggg	tgtgtntgat	240
aaantccanc	angttctcct	tggtgacctc	cccttcaaag	ttgttcgggc	cttcatacaa	300
cttctnnaan	angannancc	cancctttgtc	gagctggnat	ttgganaaca	cgtcactggt	360
ggaaactgat	cccaaagtgt	atgtcatcca	tgcctctgc	tgcttgcaaa	aaacttgctt	420
ggcncaaadc	cgactcccn	tccttgaaag	aagccnatca	cacccccctc	cctggactcc	480
nncaangact	ctnccgctnc	cccntccnng	cagggttggt	ggcannccgg	gccccntgcg	540
ttcttcagcc	agttcacnat	nttcatcagc	ccctctgcc	gctgtnttat	tccttggggg	600
ggaanccgtc	tctcccttcc	tgaannaact	ttgaccgtng	gaatagccgc	gcntcnccnt	660
acntnctggg	cgggttcaa	antccctccn	ttgcnntcn	cctcgggcc	ttctggattt	720
ncnaacttt	ttccttccc	cncctccnng	ngtttggnnt	ttcatnngg	ccccaaactc	780
gctnttgccc	antcccttgg	gggcntntan	cncctcctnt	ggtcccntng	ggcc	834

<213> Homo sapien

<223> n = A, T, C or G

cggncgcttt	cncgcgcgc	ccggtttcca	tgacnaaggc	tcccttcang	ttaaatacnn	60
cctagnaaac	attaatgggt	tgctctacta	atacatcata	cnaaccagta	agcctgccca	120
naacgccaac	tcaggccatt	cctaccaaag	gaagaaaggc	tggtctctcc	acccctgtga	180
ggaaaggcct	gccttgtaag	acaccacaat	ncggctgaat	ctnaagtctt	gtgttttact	240
aatggaaaaa	aaaaataaac	aanaggtttt	gttctcatgg	ctgcccaccg	cagcctggca	300
ctaaaaacanc	ccagcgctca	cttctgcttg	ganaaatatt	ctttgctctt	ttggacatca	360
ggcttgatgg	tatcaactgcc	acntttccac	ccagctgggc	ncccttcccc	catntttgtc	420
antganctgg	aaggcctgaa	ncttagtctc	caaaagtctc	ngcccacaag	accggccacc	480
agggggangtc	ntttncagtg	gatctgccaa	anantaccn	tatcatcnnt	gaataaaaag	540
gccctgaac	ganatgcttc	cancancctt	taagacccat	aatcctngaa	ccatggtgcc	600
cttccgggtct	gatccnaaag	gaatgttctt	gggtcccant	cctccttttg	ttncttacgt	660
tgtnttggac	ccntgctngn	atnacccaan	tganatcccc	ngaagcacc	tnccctggc	720
atttganttt	cntaaattct	ctgccctacn	nctgaaagca	cnattccctn	ggcncnnaan	780
gngaaactca	aqaaqqtctn	nqaaaaacca	cncn			814

<213> Homo sapien

<223> n = A, T, C or G



gcacgtctgct	cttctctcaaa	gttggtctcttg	ttgccataac	aaccaccata	ggtaaagcgg	60
gcgcagtggtt	cgctgaaggg	gttgtagtac	cagcgcgga	tgctctcctt	gcagagtcct	120
gtgtctggca	ggtcacgca	atgccctttg	tactggga	aatggatgcg	ctggagctcg	180
tcnaanccac	tcgtgtattt	ttcacangca	gcctcctccg	aagcntccgg	gcagttgggg	240
gtgtcgtcac	actccactaa	actgtcgatn	cancagccca	ttgctgcagc	ggaactgggt	300
gggctgacag	gtgccagaac	acactggatn	ggcctttcca	tggaaaggcc	tgggggaaat	360
cncctnancc	caaactgcct	ctcaaaggcc	accttgca	ccccgacagg	ctagaaatgc	420
actcttcttc	caaaggtag	ttgttcttgt	tgcccaagca	ncctccanca	aacaaaaanc	480
ttgcaaaatc	tgctccgtgg	gggtcatnnn	taccanggtt	ggggaaaanaa	acccggcngn	540
ganccnctt	gtttgaatgc	naaggnaata	atcctcctgt	cttgcttggg	tggaaanagca	600
caattgaact	gttaacnttg	ggcnggttc	cncnnggtg	gtctgaaact	aatcacctgc	660
actgaaaaaa	ggtangtgcc	ttccttgaat	tcccaaannt	ccctngntt	tgggtntttt	720
ctcctctncc	ctaaaaatcg	tnttcccccc	cctanggcg			760

<213> Homo sapien

<223> n = A,T,C or G

tttttttttt	tttttttttt	tttttttttt	tttttaaaaa	ccccctccat	tgaatgaaaa	60
cttcnnaaat	tgtccaaccc	cctcnnccaa	atnnccattt	cggggggggg	gttccaaacc	120
caaattaatt	ttgganttta	aattaaatnt	tnattngggg	aanaanccaa	atgtnaagaa	180
aatttaaccc	attatnaact	taaatncctn	gaaacccttg	gnttccaaaa	atttttaacc	240
cttaaatccc	tccgaaattg	ntaanggaaa	accaaattcn	cctaaggctn	tttgaagggt	300
ngatttaaac	ccccttnant	tnttttnacc	cnngnctnaa	ntatttngnt	tccggtgttt	360
tcctnttaan	cntnggtaac	tcccgntaat	gaannnccct	aanccaatta	aaccgaattt	420
tttttgaatt	ggaaattccn	ngggaattna	cgggggtttt	tcccntttgg	gggccatncc	480
cccnctttcg	gggtttgggn	ntaggttgaa	tttttnnang	ncccaaaaaa	ncccccaana	540
aaaaaaactc	caagnnttaa	ttngaattnc	ccccttccca	ggccttttgg	gaaaggnggg	600
ttntnggggg	ccnggggantt	cnttcccccn	ttncnccccc	ccccccnggt	aaanggttat	660
ngnntttggt	ttttgggccc	cttnanggac	cttcgggatn	gaaattaaat	ccccggngcg	720
gccg						724

<213> Homo sapien

<223> n = A, T, C or G







```
<210> 42
<211> 101
<212> DNA
<213> Homo sapien
```

```
<210> 43
<211> 305
<212> DNA
<213> Homo sapien
```

```
<210> 44
<211> 852
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(852)
<223> n = A,T,C or G
```

<400>	44						
acataaatat	cagagaaaag	tagtctttga	aatattttacg	tccaggagtt	ctttgtttct		60
gattattttg	tgtgtgtttt	ggtttgtgtc	caaagtattg	gcagcttcag	ttttcatttt		120
ctctccatcc	tcgggcattc	ttcccaaatt	tatataccag	tcttcgtcca	tccacacgct		180
ccagaatttc	tcttttgtag	taatatctca	tagctcggct	gagcttttca	taggtcatgc		240
tgctgttggt	cttcttttta	ccccatagct	gagccactgc	ctctgatttc	aagaacctga		300
agacgccctc	agatcggtct	tcccatttta	ttaatcctgg	gttcttgbct	gggttcaaga		360
ggatgtcgcg	gatgaattcc	cataagttag	tccctctcgg	gttggtgctt	ttggtgtggc		420
acttggcagg	ggggtcttgc	tcctttttca	tatcagggtga	ctctgcaaca	ggaagggtgac		480
tggtggttgt	catggagatc	tgagcccggc	agaaagtttt	gctgtccaac	aaatctactg		540
tgctaccata	gttggtgtca	tataaatagt	tctngtcttt	ccagggtgtc	atgatggaag		600
gctcagtttg	ttcagtcctg	acaatgacat	tgtgtgtgga	ctggaacagg	tcactactgc		660
actggccggt	ccacttcaga	tgctgcaagt	tgctgtagag	gagntgcccc	gccgtccctg		720
ccgcccgggt	gaactcctgc	aaactcatgc	tgcaaagggtg	ctcgccggtg	atgtcgaaact		780
cntggaaagg	gatacaattg	gcattccagct	ggttggtgtc	caggagggtga	tggagccact		840
cccacacctg	qt						852



<400> 45

```
<210> 46
<211> 590
<212> DNA
<213> Homo sapien
```

<400> 46

```
<210> 47
<211> 774
<212> DNA
<213> Homo sapien
```

<400> 47

acaagggggc	ataatgaagg	agtggggana	gattttaaag	aaggaaaaaa	aacgaggccc	60
tgaacagaat	tttctgnac	aacggggctt	caaaataatt	ttcttgggga	ggttcaagac	120
gcttcactgc	ttgaaactta	aatggatgtg	ggacanaatt	ttctgtaatg	accttgaggg	180
cattacagac	gggactcttg	gaggaaggat	aaacagaaaag	gggacaaagg	ctaattcccaa	240
aacatcaaag	aaaggaaggt	ggcgtcatac	ctcccagcct	acacagttct	ccagggctct	300
cctcatccct	ggaggacgac	agtggaggaa	caactgacca	tgtccccagg	ctcctgtgtg	360
ctggctcctg	gtcttcagcc	cccagctctg	gaagcccacc	ctctgctgat	cctgcgtggc	420



```

ccacactcct tgaacacaca tccccagggt atattcctgg acatggctga acctcctatt    480
cctacttccg agatgccttg ctccctgcag cctgtcaaaa tcccactcac cctccaaacc    540
acggcatggg aagcctttct gacttgcctg attactccag catcttggaa caatccctga    600
ttccccactc cttagaggca agatagggtg gttaagagta gggctggacc acttggagcc    660
aggctgctgg cttcaaattn tggctcattt acgagctatg ggaccttggg caagtnatct    720
tcacttctat gggcntcatt ttgttctacc tgcaaaatgg gggataataa tagt        774

```

<210> 48

<211> 124

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(124)

<223> n = A,T,C or G

<400> 48

```

canaaattga aattttataa aaaggcattt ttctcttata tccataaaat gatataattt    60
ttgcaantat anaaatgtgt cataaattat aatgttcctt aattacagct caacgcaact    120
tggt                                              124

```

<210> 49

<211> 147

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(147)

<223> n = A,T,C or G

<400> 49

```

gccgatgcta ctatttttatt gcaggagggtg ggggtgtttt tattattctc tcaacagctt    60
tgtggctaca ggtgggtgtct gactgcatna aaaanttttt tacgggtgat tgcaaaaatt    120
ttagggcacc catatcccaa gcantgt                    147

```

<210> 50

<211> 107

<212> DNA

<213> Homo sapien

<400> 50

```

acattaaatt aataaaagga ctgttgggggt tctgctaaaa cacatggctt gatatatattgc    60
atggttttgag gttaggagga gttaggcata tgttttggga gaggggt                    107

```

<210> 51

<211> 204

<212> DNA

<213> Homo sapien



```

<400> 51
gtcctaggaa gtctagggga cacacgactc tggggtcacg gggccgacac acttgcacgg      60
cggaaggaa aggagagaa gtgacaccgt cagggggaaa tgacagaaag gaaaatcaag      120
gccttgcaag gtcagaaagg ggactcaggg cttccaccac agccctgccc cacttggcca      180
cctccctttt gggaccagca atgt                                           204

```

```

<210> 52
<211> 491
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(491)
<223> n = A,T,C or G

```

```

<400> 52
acaaagataa catttatctt ataacaaaaa tttgatagtt ttaaagggtta gtattgtgta      60
gggtattttt caaaagacta aagagataac tcaggtaaaa agttagaaat gtataaaaca      120
ccatcagaca ggttttttaa aaacaacata ttacaaaatt agacaatcat ccttaaaaaa      180
aaaacttctt gtatcaattt cttttgttca aaatgactga cttantattt tttaaatatt      240
tcanaaacac ttcctcaaaa attttcaana tggtagcttt canatgtnc ctcagtccca      300
atgttgctca gataaataaa tctcgtgaga acttaccacc caccacaagc tttctggggc      360
atgcaacagt gtcttttctt tnccttttct tttttttttt ttacaggcac agaaactcat      420
caattttatt tggataacaa aggggtctcca aattatattg aaaaataaat ccaagttaat      480
atcactcttg t                                                         491

```

```

<210> 53
<211> 484
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(484)
<223> n = A,T,C or G

```

```

<400> 53
acataattta gcagggctaa ttaccataag atgctattta ttaanaggtn tatgatctga      60
gtattaacag ttgctgaagt ttgggtatttt tatgcagcat tttctttttg ctttgataac      120
actacagaac ccttaaggac actgaaaatt agtaagtaaa gttcagaaac attagctgct      180
caatcaaadc tctacataac actatagtaa ttaaaacggt aaaaaaaagt gttgaaatct      240
gcactagtat anaccgctcc tgtcaggata anactgcttt ggaacagaaa gggaaaaanc      300
agctttgant ttctttgtgc tgatangagg aaaggctgaa ttaccttggt gcctctccct      360
aatgattggc aggtcnggta aatnccaaaa catattccaa ctcaacactt cttttccncg      420
tancttgant ctgtgtattc caggancagg cggatggaat gggccagccc ncggtatgtc      480
cant                                                                484

```

```

<210> 54
<211> 151
<212> DNA

```



$\langle 223 \rangle$  n = A, T, C or G



&lt;400&gt; 58

acagggatat aggtttnaag ttattgtnat tgtaaaatac attgaatttt ctgtatactc	60
tgattacata catttatcct ttaaaaaaga tgtaaattctt aatttttatg ccatctatta	120
atttaccaat gagttacctt gttaatgaga agtcatgata gcactgaatt ttaactagtt	180
ttgacttcta agtttgggt	198

&lt;210&gt; 59

&lt;211&gt; 330

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 59

acaacaaatg ggttgtaggg aagtcttatac agcaaaaactg gtgatggcta ctgaaaagat	60
ccattgaaaa ttatcattaa tgatttttaa tgacaagtta tcaaaaactc actcaatttt	120
cacctgtgct agcttgctaa aatgggagtt aactctagag caaatatagt atcttctgaa	180
tacagtcaat aaatgacaaa gccagggcct acaggtgggt tccagacttt ccagaccag	240
cagaaggaat ctattttatc acatggatct cegtctgtgc tcaaaatacc taatgatatt	300
tttcgtcttt attggacttc tttgaagagt	330

&lt;210&gt; 60

&lt;211&gt; 175

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 60

accgtgggtg ctttctacat tcttgacggc tccttcacca acatctgggt ctacttcggc	60
gtcgtgggtc ctttctctt catctcctc cagctgggtg tgctcatcga ctttgcgcac	120
tcctggaacc agcgggtggc gggcaaggcc gaggagtgcg attcccgtgc ctggt	175

&lt;210&gt; 61

&lt;211&gt; 154

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 61

acccactttt tcttctgtg agcagtctgg acttctcact gctacatgat gagggtgagt	60
ggttgttgct cttcaacagt atcctccctt ttccggatct gctgagccgg acagcagtgc	120
tggactgcac agccccgggg ctccacattg ctgt	154

&lt;210&gt; 62

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 62

cgctcgagcc ctatagttag tcgtattaga	30
----------------------------------	----

&lt;210&gt; 63

&lt;211&gt; 89

&lt;212&gt; DNA

&lt;213&gt; Homo sapien



<400> 63  
 acaagtcatt tcagcaccct ttgctcttca aaactgacca tcttttatat ttaatgcttc 60  
 ctgtatgaat aaaaatggtt atgtcaagt 89

<210> 64  
 <211> 97  
 <212> DNA  
 <213> Homo sapien

<400> 64  
 accggagtaa ctgagtcggg acgctgaatc tgaatccacc aataaataaa ggttctgcag 60  
 aatcagtgc aaccaggattg gtccttggat ctggggt 97

<210> 65  
 <211> 377  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(377)  
 <223> n = A,T,C or G

<400> 65  
 acaacaanaa ntcccttctt taggccactg atggaaacct ggaacccctt tttgatggca 60  
 gcatggcgct ctaggccttg acacagcggc tgggggttgg gctntcccaa accgcacacc 120  
 ccaaccctgg tctaccaca nttctggcta tgggctgtct ctgccactga acatcagggt 180  
 tcggtcataa natgaaatcc caanggggac agaggctcagt agaggaagct caatgagaaa 240  
 ggtgctgttt gtcagccag aaaacagctg cctggcattc gccgctgaac tatgaaccg 300  
 tgggggtgaa ctaccccan gaggaatcat gcctgggcga tgcaanggtg ccaacaggag 360  
 gggcgggagg agcatgt 377

<210> 66  
 <211> 305  
 <212> DNA  
 <213> Homo sapien

<400> 66  
 acgccttttc ctcagaattc agggaagaga ctgtcgctg ccttcctccg ttgttgctg 60  
 agaaccctg tgcccttcc caccatatcc accctcgctc catctttgaa ctcaaacacg 120  
 aggaactaac tgcaccctgg tctctcccc agtccccagt tcaccctcca tccctcacct 180  
 tcctccactc taagggatat caacactgcc cagcacaggg gccctgaatt tatgtggtt 240  
 ttatatattt ttaataaga tgcactttat gtcatttttt aataaagtct gaagaattac 300  
 tgttt 305

<210> 67  
 <211> 385  
 <212> DNA  
 <213> Homo sapien



<400> 67

<210> 68

<400> 68

<210> 69

<220>

<400> 69

<210> 70

$\leq 400 > 70$



accgaaacca aattattcaa agcactgctt attacaattt tactgggtct ctatattt 477

<210> 71  
 <211> 533  
 <212> DNA  
 <213> Homo sapien  
  
 <220>  
 <221> misc\_feature  
 <222> (1)...(533)  
 <223> n = A,T,C or G

<400> 71  
 agagctatag gtacagtgtg atctcagctt tgcaaacaca ttttctacat agatagtact 60  
 aggtattaat agatatgtaa agaaagaaat cacaccatta ataatggtaa gattgggttta 120  
 tgtgatttta gtggtatttt tggcaccctt atatatgttt tccaaacttt cagcagtgat 180  
 attatttcca taacttaaaa agtgagtttg aaaaagaaaa tctccagcaa gcatctcatt 240  
 taaataaagg tttgtcatct ttaaaaatac agcaatatgt gactttttta aaaagctgtc 300  
 aaataggtgt gaccctacta ataattatta gaaatacatt taaaaacatc gagtacctca 360  
 agtcagtttg ccttgaaaaa tatcaaatat aactcttaga gaaatgtaca taaaagaatg 420  
 cttcgttaatt ttggagtang aggttccctc ctcaattttg tattttttaa aagtacatgg 480  
 taaaaaaaaa aattcacaac agtatataag gctgtaaaaat gaagaattct gcc 533

<210> 72  
 <211> 511  
 <212> DNA  
 <213> Homo sapien  
  
 <220>  
 <221> misc\_feature  
 <222> (1)...(511)  
 <223> n = A,T,C or G

<400> 72  
 tattacggaa aaacacacca cataattcaa ctancaaaga anactgcttc agggcgtgta 60  
 aaatgaaaagg cttccaggca gttatctgat taaagaacac taaaagaggg acaaggctaa 120  
 aagccgcagg atgtctacac tatancaggc gctattttggg ttggctggag gagctgtgga 180  
 aaacatggan agattgggtgc tgganatcgc cgtggctatt cctcattgtt attacanagt 240  
 gaggttctct gtgtgcccac tggtttgaaa accgttctnc aataatgata gaatagtaca 300  
 cacatgagaa ctgaaatggc ccaaaccocag aaagaaagcc caactagatc ctcagaanac 360  
 gcttctaggg acaataaccg atgaagaaaa gatggcctcc ttgtgcccc gtctgttatg 420  
 atttctctcc attgcagcna naaaccogtt cttctaagca aacncagggtg atgatggcna 480  
 aaatacacc cctcttgaag naccnggagg a 511

<210> 73  
 <211> 499  
 <212> DNA  
 <213> Homo sapien  
  
 <220>  
 <221> misc\_feature



<222> (1)...(499)

<223> n = A,T,C or G

<400> 73

cagtgccagc	actggtgcc	gtaccagtac	caataacagt	gccagtgcc	gtgccagcac	60
cagtgggtggc	ttcagtgtg	gtgccagcct	gaccgccact	ctcacatttg	ggctcttcgc	120
tggccttggg	ggagctggg	ccagcaccag	tggcagctct	gggtgcctgtg	gtttctccta	180
caagtgagat	tttagatatt	gttaatcctg	ccagtctttc	tcttcaagcc	aggggtgcac	240
ctcagaaacc	tactcaacac	agcactctag	gcagccacta	tcaatcaatt	gaagttgaca	300
ctctgcatta	aatctatttg	ccatttctga	aaaaaaaaaa	aaaaaaagg	cggccgctcg	360
antctagagg	gcccggttaa	acccgctgat	cagcctcgac	tgtgccttct	anttgccagc	420
catctgttgt	ttgccctcc	cccgtgcct	tccttgacct	tggaaagtgc	cactcccat	480
gtcctttcct	aantaaaat					499

<210> 74

<211> 537

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(537)

<223> n = A,T,C or G

<400> 74

tttcatagga	gaacacactg	aggagatact	tgaagaat	ggattcagcc	gcgaagagat	60
ttatcagctt	aactcagata	aaatcattga	aagtaataag	gtaaaagcta	gtctctaact	120
tccaggccca	cggctcaagt	gaatttgaat	actgcattta	cagtgtagag	taacacataa	180
cattgtatgc	atggaaacat	ggaggaacag	tattacagtg	tcctaccact	ctaatacaaga	240
aaagaattac	agactctgat	tctacagtga	tgattgaatt	ctaaaaatgg	taatcattag	300
ggcttttgat	ttataaanact	ttgggtactt	atactaaatt	atggtagtta	tactgccttc	360
cagtttgctt	gatatatattg	ttgatattaa	gattccttgac	ttatatatttg	aatgggttct	420
actgaaaaan	gaatgatata	ttcttgaaga	catcgatata	cattttattta	cactcttgat	480
tctacaatgt	agaaaatgaa	ggaaatgccc	caaattgtat	gggtataaaa	gtcccg	537

<210> 75

<211> 467

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(467)

<223> n = A,T,C or G

<400> 75

caaanacaat	tgttcaaaag	atgcaaata	tacactactg	ctgcagctca	caaacacctc	60
tgcattattac	acgtacctcc	tcctgtctct	caagtagtgt	ggctctat	gccatcatca	120
cctgtgtct	gcttagaaga	acggctttct	gctgcaangg	agagaaatca	taacagacgg	180
tggcacaagg	aggccatctt	ttcctcatcg	gttattgtcc	ctagaagcgt	cttctgagga	240
tctagttggg	ctttctttct	gggtttgggc	catttcantt	ctcatgtgtg	tactattcta	300



```

tcattattgt ataacggttt tcaaaccngt gggcacncag agaacctcac tctgtaataa 360
caatgaggaa tagccacggt gatctccagc accaaatctc tccatgttnt tccagagctc 420
ctccagccaa cccaaatagc cgctgctatn gtgtagaaca tccttgn 467

```

```

<210> 76
<211> 400
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(400)
<223> n = A,T,C or G

```

```

<400> 76
aagctgacag cattcgggcc gagatgtctc gctccgtggc cttagctgtg ctgcgcctac 60
tctctctttc tggcctggag gctatccagc gtactccaaa gattcaggtt tactcacgtc 120
atccagcaga gaatggaaag tcaaatttcc tgaattgcta tgtgtctggg tttcatccat 180
ccgacattga agttgactta ctgaagaatg gagagagaat tgaaaaagtg gagcattcag 240
acttgtcttt cagcaaggac tgggtctttct atctcttgta ctacactgaa ttcaccccca 300
ctgaaaaaga tgagtatgcc tgccgtgtga accatgtgac tttgtcacag cccaagatng 360
ttnagtggga tcganacatg taagcagcan catgggaggt 400

```

```

<210> 77
<211> 248
<212> DNA
<213> Homo sapien

```

```

<400> 77
ctggagtgcc ttggtgtttc aagcccctgc aggaagcaga atgcaccttc tgaggcacct 60
ccagctgccc cggcggggga tgcgaggctc ggagcaccct tgcccggctg tgattgctgc 120
caggcactgt tcatctcagc ttttctgtcc ctttgetccc ggcaagcgct tctgctgaaa 180
gttcatatct ggagcctgat gtcttaacga ataaaggctc catgctccac ccgaaaaaaa 240
aaaaaaaaa 248

```

```

<210> 78
<211> 201
<212> DNA
<213> Homo sapien

```

```

<400> 78
actagtccag tgtggtggaa ttccattgtg ttgggcccac cacaatggct acctttaaca 60
tcaccagac ccgcacctgc ccgtgccccca cgctgctgct aacgacagta tgatgcttac 120
tctgctactc ggaaactatt tttatgtaat taatgtatgc tttcttgttt ataaatgcct 180
gatttaaaaa aaaaaaaaaa a 201

```

```

<210> 79
<211> 552
<212> DNA
<213> Homo sapien

```



<220>  
 <221> misc\_feature  
 <222> (1)...(552)  
 <223> n = A,T,C or G

<400> 79  
 tccttttgtt aggtttttga gacaacccta gacctaaact gtgtcacaga cttctgaatg 60  
 tttaggcagt gctagtaatt tcctcgtaat gattctgtta ttactttcct attctttatt 120  
 cctctttctt ctgaagatta atgaagttga aaattgaggt ggataaatac aaaaaggtag 180  
 tgtgatagta taagtatcta agtgcagatg aaagtgtgtt atatatatcc attcaaaatt 240  
 atgcaagtta gtaattactc agggtttaact aaattacttt aatatgctgt tgaacctact 300  
 ctgttccttg gctagaaaaa attataaaca ggactttgtt agtttgggaa gccaaattga 360  
 taatattcta tgttctaaaa gttgggctat acataaanta tnaagaaata tggaatttta 420  
 ttcccaggaa tatgggggttc atttatgaat antaccggg anagaagttt tgantnaaac 480  
 cngtttttgt taatacgta atatgtcctn aatnaacaag gcntgactta tttccaaaaa 540  
 aaaaaaaaaa aa 552

<210> 80  
 <211> 476  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(476)  
 <223> n = A,T,C or G

<400> 80  
 acagggattt gagatgctaa ggccccagag atcgtttgat ccaaccctct tattttcaga 60  
 ggggaaaatg gggcctagaa gttacagagc atctagctgg tgcgctggca cccctggcct 120  
 cacacagact cccgagtagc tgggactaca ggcacacagt cactgaagca ggccctgttt 180  
 gcaattcacg ttgccacctc caacttaaac attcttcata tgtgatgtcc ttagtcacta 240  
 aggttaaaact ttcccacca gaaaaggcaa cttagataaa atcttagagt actttcatac 300  
 tcttctaagt cctcttcag cctcactttg agtcctcctt ggggggttgat aggaantntc 360  
 tcttggtttt ctcaataaaa tctctatcca tctcatgttt aatttggtac gcntaaaaat 420  
 gctgaaaaaa ttaaaatgtt ctggtttcnc tttaaaaaaa aaaaaaaaaa aaaaaa 476

<210> 81  
 <211> 232  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(232)  
 <223> n = A,T,C or G

<400> 81  
 tttttttttg tatgcntcn ctgtggngtt attgttgctg ccaccctgga ggagcccagt 60  
 ttcttctgta tctttctttt ctgggggagc ttcctggctc tgccctcca tcccagcct 120  
 ctcacccca tcttgcaactt ttgctagggg tggaggcgct ttcctggtag cccctcagag 180



actcagtcag cggaataag tcctaggggt ggggggtgtg gcaagccggc ct 232

<210> 82  
 <211> 383  
 <212> DNA  
 <213> Homo sapien  
 <220>  
 <221> misc\_feature  
 <222> (1)...(383)  
 <223> n = A,T,C or G

<400> 82  
 aggcgggagc agaagctaaa gccaaagccc aagaagagtg gcagtgccag cactggtgcc 60  
 agtaccagta ccaataacat gccagtgcc gtgccagcac cagtgggtggc ttcagtgctg 120  
 gtgccagcct gaccgccact ctcacatttg ggcctttgc tggccttggg ggagctggtg 180  
 ccagcaccag tggcagctct ggtgcctgtg gtttctccta caagtgagat tttagatatt 240  
 gttaatcctg ccagtccttc tcttcaagcc aggggtgcac ctcagaaacc tactcaacac 300  
 agcactctng gcagccacta tcaatcaatt gaagttgaca ctctgcatta aatctatttg 360  
 ccatttcaaa aaaaaaaaaa aaa 383

<210> 83  
 <211> 494  
 <212> DNA  
 <213> Homo sapien  
 <220>  
 <221> misc\_feature  
 <222> (1)...(494)  
 <223> n = A,T,C or G

<400> 83  
 accgaattgg gaccgtggc ttataagcga tcatgtctc cagtattacc tcaacgagca 60  
 gggagatcga gtctatacgc tgaagaaatt tgaccgatg ggacaacaga cctgctcagc 120  
 ccatcctgct cggttctccc cagatgacaa atactctcga caccgaatca ccatcaagaa 180  
 acgcttcaag gtgctcatga cccagcaacc gcgcctgtc ctctgagggt ccttaaactg 240  
 atgtcttttc tgccacctgt taccctcgg agactccgta accaaactct tcggactgtg 300  
 agccctgatg cttttttgcc agccatactc tttggcntcc agtctctcgt ggcgattgat 360  
 tatgcttgtg tgaggcaatc atgggtggcat caccatnaa gggaacacat ttganttttt 420  
 tttcncatat tttaaattac naccagaata ntccagaata aatgaattga aaaactctta 480  
 aaaaaaaaaa aaaa 494

<210> 84  
 <211> 380  
 <212> DNA  
 <213> Homo sapien  
 <220>  
 <221> misc\_feature  
 <222> (1)...(380)  
 <223> n = A,T,C or G



&lt;400&gt; 84

```

gctggtagcc tatggcgtgg ccacggangg gctcctgagg cacgggacag tgacttccca      60
agtatcctgc gccgcgtctt ctaccgtccc tacctgcaga tcttcgggca gattccccag      120
gaggacatgg acgtggccct catggagcac agcaactgct cgtcggagcc cggcttctgg      180
gcacaccctc ctggggccca ggcgggcacc tgcgtctccc agtatgcaa ctggctggtg      240
gtgctgctcc tcgtcatctt cctgctcgtg gccaacatcc tgctggtcac ttgctcattg      300
ccatgttcag ttacacattc ggcaaagtac agggcaacag cnatctctac tgggaaggcc      360
agcgttnccg cctcatccgg

```

&lt;210&gt; 85

&lt;211&gt; 481

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(481)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 85

```

gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggcctctcgc ttcataccgc      60
tnccatcgtc atactgtagg tttgccacca cctcctgcat cttggggcgg ctaatatcca      120
ggaaactctc aatcaagtca ccgtcnatna aacctgtggc tggttctgtc ttccgctcgg      180
tgtgaaagga tctccagaag gagtgtctga tcttccccac acttttgatg actttattga      240
gtcgattctg catgtccagc aggaggttgt accagctctc tgacagtgag gtcaccagcc      300
ctatcatgcc nttgaacgtg ccgaagaaca ccgagccttg tgtggggggg gnagtctcac      360
ccagattctg cattaccaga nagccgtggc aaaaganatt gacaactcgc ccaggngaa      420
aaagaacacc tcctggaagt gctngccgct cctcgtccnt tgggtggngc gcntnccttt      480
t

```

&lt;210&gt; 86

&lt;211&gt; 472

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(472)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 86

```

aacatcttcc tgtataatgc tgtgtaatat cgatccgatn ttgtctgctg agaattcatt      60
acttggaana gcaacttnaa gcctggacac tgggtattaaa attcacaata tgcaaacatt      120
taaacagtgt gtcaatctgc tcccttactt tgtcatcacc agtctgggaa taagggtatg      180
ccctattcac acctgttaaa agggcgctaa gcatttttga ttcaacatct ttttttttga      240
cacaagtccg aaaaaagcaa aagtaaacag ttnttaattt gttagccaat tcactttctt      300
catgggacag agccatttga tttaaaaagc aaattgcata atattgagct ttgggagctg      360
atatntgagc ggaagantag cctttctact tcaccagaca caactccttt catattggga      420
tgttnacnaa agttatgtct cttacagatg ggatgctttt gtggcaattc tg      472

```



<210> 87  
 <211> 413  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(413)  
 <223> n = A,T,C or G

<400> 87  
 agaaaccagt atctctnaaa acaacctctc ataccttggtg gacctaatTT tgtgtgcgtg 60  
 tgtgtgtgcg cgcataattat atagacaggc acatctTTTT tacttttgta aaagcttatg 120  
 cctctttggt atctatatct gtgaaagttt taatgatctg ccataatgtc ttggggacct 180  
 ttgtcttctg tgtaaatggt actagagaaa acacctatnt tatgagtcaa tctagttngt 240  
 tttattcgac atgaaggaaa tttccagatn acaacactna caaactctcc cttgactagg 300  
 ggggacaaaag aaaagcnaaa ctgaacatna gaaacaattn cctgggtgaga aattncataa 360  
 acagaaattg ggtngtatat tgaaanannng catcattnaa acgtTTTTTT ttt 413

<210> 88  
 <211> 448  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(448)  
 <223> n = A,T,C or G

<400> 88  
 cgcagcgggt cctctctatc tagctccagc ctctcgctg cccactccc cgcgtcccgc 60  
 gtcctagccn accatggccg ggccccctgcg cgccccgcgtg ctctgtgctg ccacctggc 120  
 cgtggccctg gccgtgagcc ccgcggcccg ctccagtcct ggcaagccgc cgcgcctggt 180  
 gggaggccca tggacccccgc gtggaagaag aagggtgtgcg gcgtgcactg gactttgccc 240  
 tcggcnanta caacaaaccc gcaacnactt ttaccnagcn cgcgctgcag gttgtgccgc 300  
 cccaancaa ttgttactng gggtaantaa ttcttggaag ttgaacctgg gccaaacnng 360  
 tttaccagaa ccnagccaat tngaacaatt nccccctcat aacagcccct tttaaaaagg 420  
 gaancantcc tgntcttttc caaatTTT 448

<210> 89  
 <211> 463  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(463)  
 <223> n = A,T,C or G

<400> 89  
 gaattttgtg cactggccac tgtgatggaa ccattgggcc aggatgcttt gagtttatca 60



```
<210> 90
<211> 400
<212> DNA
<213> Homo sapien
```

<400>	90						
ttgaa	ggtctntnt	actgtcggac	tgttcanca	ccaactctac	aagttgctgt		60
actca	ctgtctgtaa	gcntnttaac	ccagactgta	tcttcataaa	tagaacaaat		120
accag	tcacatcttc	taggaccttt	ttggattcag	ttagtataag	ctcttccact		180
cgta	agacttcate	tggtaaagtc	ttagttttg	tagaaaggaa	tttaattgct		240
cttaa	caatgtcttc	tccttgaagt	atttggtgta	acaaccacc	tnaagtcct		300
catcc	attttaaata	tacttaatag	ggcattggtg	cactaggtta	aattctgcaa		360
atctg	tctgcaaaag	ttgcgttagt	atatctgcc				400

```
<210> 91
<211> 480
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1) ... (480)
<223> n = A,T,C or G
```

```
<210> 92
<211> 477
<212> DNA
<213> Homo sapien
```



<220>  
 <221> misc\_feature  
 <222> (1)...(477)  
 <223> n = A,T,C or G

<400> 92  
 atacagccca natcccacca cgaagatgcg cttgttgact gagaacctga tgcggtcact 60  
 ggtcccgtg tagccccagc gactctccac ctgctggaag cggttgatgc tgcactcctt 120  
 cccacgcagg cagcagcggg gccggccaat gaactccact cgtggcttgg gggtgacggg 180  
 taantgcagg aagaggctga ccacctcgcg gtccaccagg atgcccgact gtgcgggacc 240  
 tgcagcgaaa ctctctgatg gtcattgagcg ggaagcgaat gangcccagg gccttgccca 300  
 gaaccttccg cctgttctct ggcgtcacct gcagctgctg ccgctnacac tcggcctcgg 360  
 accagcggac aaacggcggt gaacagccgc acctcacgga tgcccantgt gtcgcgctcc 420  
 aggaacggcn ccagcgtgtc caggtcaatg tcggtgaanc ctccgcgggt aatggcg 477

<210> 93  
 <211> 377  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(377)  
 <223> n = A,T,C or G

<400> 93  
 gaacggctgg accttgctc gcattgtgct gctggcagga ataccttggc aagcagctcc 60  
 agtccgagca gccccagacc gctgccgccc gaagctaagc ctgcctctgg ccttcccctc 120  
 cgcctcaatg cagaaccant agtgggagca ctgtgttttag agttaagagt gaacactgtn 180  
 tgattttact tgggaatttc ctctgttata tagcttttcc caatgctaata ttccaaacaa 240  
 caacaacaaa ataacatggt tgccctgttna gttgtataaaa agtangtgat tctgtatnta 300  
 aagaaaatat tactgttaca tatactgctt gcaantttctg tattttattgg tnctctggaa 360  
 ataaatatat tattaata 377

<210> 94  
 <211> 495  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(495)  
 <223> n = A,T,C or G

<400> 94  
 ccctttgagg ggtaggggc cagttcccag tggaagaaac aggccaggag aantgcgtgc 60  
 cgagctgang cagatttccc acagtgacct cagagccctg ggctatagtc tctgacctct 120  
 ccaaggaaa accaccttct ggggacatgg gctggagggc aggacctaga ggcaccaagg 180  
 gaaggcccca ttccggggct gttccccgag gaggaaggga aggggctctg tgtgcccccc 240  
 acgaggaana ggccctgant cctgggatca nacaccctt cacgtgtatc cccacacaaa 300



```

tgcaagctca ccaaggtccc ctctcagtc cttccctaca ccctgaacgg nactggccc 360
acaccacccc agancancca cccgccatgg ggaatgtnc tcaaggaatcg cngggcaacg 420
tggaactctng tcccnaagg gggcagaatc tccaatagan gganngaacc cttgctnana 480
aaaaaaaaa aaaaa 495

```

```

<210> 95
<211> 472
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(472)
<223> n = A,T,C or G

```

```

<400> 95
ggttacttgg tttcattgcc accacttagt ggatgtcatt tagaaccatt ttgtctgctc 60
cctctggaag ccttgcgcag agcggacttt gtaattgttg gagaataact gctgaatttt 120
tagctgtttt gagttgattc gcaccactgc accacaactc aatatgaaaa ctatttnact 180
tatttattat cttgtgaaaa gtatacaatg aaaattttgt tcatactgta tttatcaagt 240
atgatgaaaa gcaatagata tatattcttt tattatgttn aattatgatt gccattatta 300
atcggcaaaa tgtggagtg atgttctttt cacagtaata tatgcctttt gtaacttcac 360
ttggttattt tattgtaaat gaattacaaa attcttaatt taagaaaatg gtangttata 420
tttanttcan taatttcttt ccttgtttac gttaattttg aaaagaatgc at 472

```

```

<210> 96
<211> 476
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(476)
<223> n = A,T,C or G

```

```

<400> 96
ctgaagcatt tcttcaaact tntctacttt tgtcattgat acctgtagta agttgacaat 60
gtggtgaaat ttcaaaatta tatgtaactt ctactagttt tacttttctcc cccaagtctt 120
ttttaactca tgattttttac acacacaatc cagaacttat tatatagcct ctaagtcttt 180
attcttcaca gtagatgatg aaagagtcct ccagtgtctt gngcanaatg ttctagntat 240
agctggatac atacngtggg agttctataa actcatacct cagtgggact naaccaaaat 300
tgtgttagtc tcaattccta ccacactgag ggagcctccc aaatcactat attcttatct 360
gcaggtaact ctcacagaaa acngacaggg caggcttgca tgaaaaagtn acatctgcgt 420
taciaagtct atcttcctca nangtctgtt aaggaacaat ttaatcttct agcttt 476

```

```

<210> 97
<211> 479
<212> DNA
<213> Homo sapien

```

```

<220>

```



<221> misc\_feature  
 <222> (1)...(479)  
 <223> n = A,T,C or G

<400> 97

actctttcta	atgctgatat	gatcttgagt	ataagaatgc	atatgtcact	agaatggata	60
aaataatgct	gcaaaactta	tggtcttatg	caaaatggaa	cgctaata	acacagctta	120
caatcgcaaa	tcaaaactca	caagtgtctc	tctgttgtag	atttagtgta	ataagactta	180
gattgtgctc	cttcggatat	gattgtttct	canatcttgg	gcaatnttcc	ttagtcaa	240
caggctacta	gaattctgtt	attggatatn	tgagagcatg	aaatttttaa	naatacactt	300
gtgattatna	aattaatcac	aaatttcact	tatacctgct	atcagcagct	agaaaaacat	360
ntnnttttta	natcaaagta	ttttgtgttt	ggaantgtnn	aaatgaaatc	tgaatgtggg	420
ttcnatctta	ttttttcccn	gacnactant	tnctttttta	gggnctattc	tgancatc	479

<210> 98  
 <211> 461  
 <212> DNA  
 <213> Homo sapien

<400> 98

agtgacttgt	cctccaacaa	aacccttga	tcaagtttgt	ggcactgaca	atcagaccta	60
tgctagttcc	tgcatctat	tcgctactaa	atgcagactg	gaggggacca	aaaaggggca	120
tcaactccag	ctggattatt	ttggagcctg	caaactcatt	cctacttgta	cggactttga	180
agtgattcag	tttctctac	ggatgagaga	ctggctcaag	aatatcctca	tcagacttta	240
tgaagccact	ctgaacacgc	tggttatcta	gatgagaaca	gagaaataaa	gtcagaaaat	300
ttacctggag	aaaagaggct	ttggctgggg	accatcccat	tgaaccttct	cttaaggact	360
ttaagaaaaa	ctaccacatg	ttgtgtatcc	tggtgcgggc	cgtttatgaa	ctgaccaccc	420
tttgaataaa	tcttgacgct	cctgaacttg	ctcctctgcg	a		461

<210> 99  
 <211> 171  
 <212> DNA  
 <213> Homo sapien

<400> 99

gtggccgcgc	gcaggtgttt	cctcgtagcg	cagggccccc	tcccttcccc	aggcgctccct	60
cggcgccctct	gcgggcccga	ggaggagcgg	ctggcggttg	gggggagtgt	gaccacccct	120
cggtgagaaa	agccttctct	agcgatctga	gaggcggtgc	ttgggggtac	c	171

<210> 100  
 <211> 269  
 <212> DNA  
 <213> Homo sapien

<400> 100

cggccgcaag	tgcaactcca	gctggggccg	tgcgagcgaa	gattctgcca	gcagttggtc	60
cgactgcgac	gacggcggcg	gcgacagtcg	caggtgcagc	gcggggcgct	ggggtcttgc	120
aaggctgagc	tgacgccgca	gaggtcgtgt	cacgtcccac	gaccttgacg	ccgtcgggga	180
cagccggaac	agagcccggt	gaagcgggag	gcctcgggga	gccctcggg	aagggcggcc	240
cgagagatac	gcaggtgcag	gtggcccgcc				269



<210> 101  
 <211> 405  
 <212> DNA  
 <213> Homo sapien

<400> 101  
 tttttttttt ttttggaaac tactgcgagc acagcaggtc agcaacaagt ttattttgca 60  
 gctagcaagg taacagggta gggcatgggt acatgttcag gtcaacttcc ttgtcgtgg 120  
 ttgattgggt tgtctttatg ggggcggggg ggggtagggg aaacgaagca aataacatgg 180  
 agtgggtgca ccctccctgt agaacctggg tacaagctt ggggcagttc acctggtctg 240  
 tgaccgtcat tttcttgaca tcaatgttat tagaagtcag gatattcttt agagagtcca 300  
 ctgttctgga gggagattag gggttcttgc caaatccaac aaaatccact gaaaaagttg 360  
 gatgatcagt acgaataccg aggcataattc tcatatcggt ggcca 405

<210> 102  
 <211> 470  
 <212> DNA  
 <213> Homo sapien

<400> 102  
 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 60  
 ggcacttaat ccatttttat ttcaaaatgt ctacaaatth aatcccatta tacggatttt 120  
 tcaaaatcta aattattcaa attagccaaa tccttaccaa ataataccca aaaatcaaaa 180  
 atatacttct ttcagcaaac ttgttacata aattaaaaaa atatatacgg ctggtgtttt 240  
 caaagtacaa ttatcttaac actgcaaaca ttttaaggaa ctaaaataaa aaaaaacact 300  
 ccgcaaagggt taaagggaac aacaaattct tttacaacac cattataaaa atcatatctc 360  
 aaatcttagg ggaatatata cttcacacgg gatcttaact ttactcact ttgtttattt 420  
 ttttaaacca ttgtttgggc ccaacacaat ggaatcccc ctggactagt 470

<210> 103  
 <211> 581  
 <212> DNA  
 <213> Homo sapien

<400> 103  
 tttttttttt ttttttttga cccccctctt ataaaaaaca agttaccatt ttattttact 60  
 tacacatatt tatttttataa ttggtattag atattcaaaa ggcagctttt aaaatcaaac 120  
 taaatggaaa ctgccttaga tacataattc ttaggaatta gcttaaaatc tgcctaaagt 180  
 gaaaaatctt tctagctctt ttgactgtaa atttttgact cttgtaaaac atccaaattc 240  
 atttttcttg tctttaaaat tatctaattc ttccattttt tccctattcc aagtcaattt 300  
 gcttctctag cctcatttcc tagctcttat ctactattag taagtggctt ttttcctaaa 360  
 agggaaaaaca ggaagagaaa tggcacacaa aacaaacatt ttatattcat atttctacct 420  
 acgttaataa aatagcattt tgtgaagcca gctcaaaaga aggcttagat ccttttatgt 480  
 ccatttttagt cactaaacga tatcaaagtg ccagaatgca aaaggtttgt gaacatttat 540  
 tcaaaagcta atataagata tttcacatac tcatctttct g 581

<210> 104  
 <211> 578  
 <212> DNA  
 <213> Homo sapien



&lt;400&gt; 104

tttttttttt	tttttttttt	tttttctctt	cttttttttt	gaaatgagga	tcgagttttt	60
cactctctag	atagggcatg	aagaaaactc	atctttccag	ctttaaaata	acaatcaa	120
ctcttatgct	atatcatatt	ttaagttaaa	ctaattgagtc	actggcttat	cttctcctga	180
aggaaatctg	ttcattcttc	tcattcatat	agttatatca	agtactacct	tgcattattga	240
gagggtttttc	ttctctat	acacatatat	ttccatgtga	atttgtatca	aacctttatt	300
ttcatgcaaa	ctagaaaata	atgtttcttt	tgcataagag	aagagaacaa	tatagcatta	360
caaaactgct	caaattgttt	gttaagttat	ccattataat	tagttggcag	gagctaatac	420
aaatcacatt	tacgacagca	ataataaaac	tgaagtacca	gttaaataatc	caaaataatt	480
aaaggaacat	ttttagcctg	ggtataatta	gctaattcac	tttacaagca	tttattagaa	540
tgaattcaca	tgttattatt	cctagcccaa	cacaatgg			578

&lt;210&gt; 105

&lt;211&gt; 538

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 105

tttttttttt	tttttcagta	ataatcagaa	caatatttat	ttttatattt	aaaattcata	60
gaaaagtgcc	ttacatttaa	taaaagtttg	tttctcaaag	tgatcagagg	aattagatat	120
gtcttgaaca	ccaatattaa	tttgaggaaa	atacaccaaa	atacattaag	taaattattt	180
aagatcatag	agcttgtaag	tgaaaagata	aaatttgacc	tcagaaactc	tgagcattaa	240
aatccacta	ttagcaata	aattactatg	gacttcttgc	tttaattttg	tgatgaatat	300
ggggtgtcac	tggtaaacca	acacattctg	aaggatacat	tacttagtga	tagattctta	360
tgtactttgc	taatacgtgg	atatgagttg	acaagtttct	ctttcttcaa	tcttttaagg	420
ggcgagaaat	gaggaagaaa	agaaaaggat	tacgcatact	gttctttcta	tggaaggatt	480
agatatgttt	cctttgccaa	tattaaaaaa	ataataatgt	ttactactag	tgaaaccc	538

&lt;210&gt; 106

&lt;211&gt; 473

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 106

tttttttttt	tttttttagtc	aagtttctat	ttttattata	attaaagtct	tggtcatttc	60
atttattagc	tctgcaactt	acatatttaa	attaaagaaa	cgtttttagac	aactgtacaa	120
tttataaatg	taagggtgcca	ttattgagta	atatattcct	ccaagagtgg	atgtgtccct	180
tctcccacca	actaatgaac	agcaacatta	gttttaatttt	attagtagat	atacactgct	240
gcaaacgcta	attctcttct	ccatcccat	gtgatattgt	gtatatgtgt	gagttggtag	300
aatgcatcac	aatctacaat	caacagcaag	atgaagctag	gctgggcttt	cggtgaaaat	360
agactgtgtc	tgtctgaatc	aaatgatctg	acctatcctc	ggtggcaaga	actcttcgaa	420
ccgcttcctc	aaaggcgctg	ccacatttgt	ggctctttgc	acttgtttca	aaa	473

&lt;210&gt; 107

&lt;211&gt; 1621

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 107

cgccatggca	ctgcagggca	tctcgggtcat	ggagctgtcc	ggcctggccc	cgggcccggtt	60
ctgtgctatg	gtcctggctg	acttcggggc	gcgtgtggta	cgctgggacc	ggcccggctc	120



```

ccgctacgac gtgagccgct tgggcccgggg caagcgcctcg ctagtgctgg acctgaagca 180
gccgcgggga gccgccgtgc tgcggcgctct gtgcaagcgg tcggatgtgc tgctggagcc 240
cttccgcccgc ggtgtcatgg agaaactcca gctggggcca gagattctgc agcgggaaaa 300
tccaaggctt atttatgcca ggctgagtggt atttggccag tcaggaagct tctgccggtt 360
agctggccac gatatcaact atttggcttt gtcaggtgtt ctctcaaaaa ttggcagaag 420
tggtgagaat ccgtatgccc cgctgaatct cctggctgac tttgctggtg gtggccttat 480
gtgtgactg ggcattataa tggctctttt tgaccgcaca cgcactgaca agggtcaggt 540
cattgatgca aatatggtgg aaggaacagc atatttaagt tcttttctgt ggaaaactca 600
gaaatcgagt ctgtgggaag cacctcgagg acagaacatg ttggatggtg gagcaccttt 660
ctatacgact tacaggacag cagatgggga attcatggct gttggagcaa tagaacccca 720
gttctacgag ctgctgatca aaggacttgg actaaagtct gatgaacttc ccaatcagat 780
gagcatggat gattggccag aaatgaagaa gaagtttgca gatgtatttg caaagaagac 840
gaaggcagag tgggtgcaaa tctttgacgg cacagatgcc tgtgtgactc cggttctgac 900
ttttgaggag gttgttcac atgatcacia caaggaacgg ggctcgttta tcaccagtga 960
ggagcaggac gtgagcccc gccctgcacc tctgctgtta aacaccccag ccatcccttc 1020
tttcaaaagg gatcctttca taggagaaca cactgaggag atacttgaag aatttggatt 1080
cagccgcgaa gagatttatc agcttaactc agataaaatc attgaaagta ataaggtaaa 1140
agctagtctc taacttccag gccacggct caagtgaatt tgaatactgc atttacagt 1200
tagagtaaca cataacattg tatgcatgga aacatggagg aacagtatta cagtgtccta 1260
ccactctaata caagaaaaga attacagact ctgattctac agtgatgatt gaattctaaa 1320
aatggttatc attagggctt ttgatttata aaactttggg tacttatact aaattatggt 1380
agttattctg ccttccagtt tgcttgatat atttgttgat attaagattc ttgacttata 1440
ttttgaatgg gttctagtga aaaaggaatg atatatctt gaagacatcg atatacat 1500
atttacactc ttgattctac aatgtagaaa atgaggaaat gccacaaatt gtatggtgat 1560
aaaagtcacg tgaaacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1620
a 1621

```

<210> 108

<211> 382

<212> PRT

<213> Homo sapien

<400> 108

```

Met Ala Leu Gln Gly Ile Ser Val Met Glu Leu Ser Gly Leu Ala Pro
1          5          10          15
Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val Val
20          25          30
Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly Arg
35          40          45
Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala Ala
50          55          60
Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro Phe
65          70          75          80
Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu Gln
85          90          95
Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly Gln
100         105         110
Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu Ala
115         120         125
Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro Tyr
130         135         140

```



Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Gly Leu Met Cys  
 145 150 155 160  
 Ala Leu Gly Ile Ile Met Ala Leu Phe Asp Arg Thr Arg Thr Asp Lys  
 165 170 175  
 Gly Gln Val Ile Asp Ala Asn Met Val Glu Gly Thr Ala Tyr Leu Ser  
 180 185 190  
 Ser Phe Leu Trp Lys Thr Gln Lys Ser Ser Leu Trp Glu Ala Pro Arg  
 195 200 205  
 Gly Gln Asn Met Leu Asp Gly Gly Ala Pro Phe Tyr Thr Thr Tyr Arg  
 210 215 220  
 Thr Ala Asp Gly Glu Phe Met Ala Val Gly Ala Ile Glu Pro Gln Phe  
 225 230 235 240  
 Tyr Glu Leu Leu Ile Lys Gly Leu Gly Leu Lys Ser Asp Glu Leu Pro  
 245 250 255  
 Asn Gln Met Ser Met Asp Asp Trp Pro Glu Met Lys Lys Lys Phe Ala  
 260 265 270  
 Asp Val Phe Ala Lys Lys Thr Lys Ala Glu Trp Cys Gln Ile Phe Asp  
 275 280 285  
 Gly Thr Asp Ala Cys Val Thr Pro Val Leu Thr Phe Glu Glu Val Val  
 290 295 300  
 His His Asp His Asn Lys Glu Arg Gly Ser Phe Ile Thr Ser Glu Glu  
 305 310 315 320  
 Gln Asp Val Ser Pro Arg Pro Ala Pro Leu Leu Leu Asn Thr Pro Ala  
 325 330 335  
 Ile Pro Ser Phe Lys Arg Asp Pro Phe Ile Gly Glu His Thr Glu Glu  
 340 345 350  
 Ile Leu Glu Glu Phe Gly Phe Ser Arg Glu Glu Ile Tyr Gln Leu Asn  
 355 360 365  
 Ser Asp Lys Ile Ile Glu Ser Asn Lys Val Lys Ala Ser Leu  
 370 375 380

<210> 109

<211> 1524

<212> DNA

<213> Homo sapien

<400> 109

```

ggcacgaggc tgcgccaggc cctgagcgga ggcgggggca gcctcgccag cggggggccc 60
gggcctggcc atgcctcact gagccagcgc ctgcgcctct acctcgccga cagctggaac 120
cagtgcgacc tagtggtctt cactgcttc ctccctggcg tgggctgccg gctgaccccg 180
ggtttgtacc acctgggccc cactgtcttc tgcctcgact tcatggtttt cacggtgcgg 240
ctgcttcaca tcttcacggc caacaaacag ctggggccca agatcgctcat cgtgagcaag 300
atgatgaagg acgtgttctt ctccctcttc ttccctcgcg tgtggctggc agcctatggc 360
gtggccacgg aggggctcct gaggccacgg gacagtgcac tcccaagtat cctgcgccgc 420
gtcttctacc gtccctacct gcagatcttc gggcagattc cccaggagga catggacgtg 480
gccctcatgg agcacagcaa ctgctcgtcg gagcccggct tctgggcaca ccctcctggg 540
gccaggcggg gcacctgcgt ctcccagtat gccaactggc tgggtgggtgt gctcctcgtc 600
atcttctctg tcgtggccaa catcctgctg gtcaacttgc tcattgccat gttcagttac 660
acattcggca aagtacaggg caacagcgat ctctactgga aggcgcagcg ttaccgcctc 720
atccgggaat tccactctcg gcccgcgctg gccccgccct ttatcgctcat ctcccacttg 780
cgctcctgct tcaggcaatt gtgcaggcga ccccgagacc ccagccgctc ctcccgggcc 840

```



ctcgagcatt	tccgggttta	ccttttctaag	gaagccgagc	ggaagctgct	aacgtgggaa	900
tcggtgcata	aggagaactt	tctgctggca	cgcgctaggg	acaagcgga	gagcgactcc	960
gagcgtctga	agcgcacgtc	ccagaaggtg	gacttggcac	tgaacagct	gggacacatc	1020
cgcgagtacg	aacagcgcc	gaaagtgtg	gagcgggagg	tccagcagtg	tagccgcgtc	1080
ctgggggtggg	tggccgaggc	cctgagccgc	tctgccttgc	tgcccccagg	tgggcccga	1140
ccccctgacc	tgcttgggtc	caaagactga	gccctgctgg	cggacttcaa	ggagaagccc	1200
ccacagggga	ttttgtcct	agagtaaggc	tcatctgggc	ctcgcccccc	gcacctggtg	1260
gccttgtcct	tgaggtgagc	cccatgtcca	tctggggccac	tgtcaggacc	acctttggga	1320
gtgtcatcct	tacaaaccac	agcatgccc	gctcctccca	gaaccagtcc	cagcctggga	1380
ggatcaaggc	ctggatcccg	ggccgttatc	catctggagg	ctgcagggtc	cttggggtaa	1440
cagggaccac	agaccctca	ccactcacag	attcctcaca	ctggggaaat	aaagccattt	1500
cagaggaaaa	aaaaaaaaaa	aaaa				1524

&lt;210&gt; 110

&lt;211&gt; 3410

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 110

gggaaccagc	ctgcacgcgc	tggctccggg	tgacagccgc	gcgcctcggc	caggatctga	60
gtgatgagac	gtgtccccac	tgaggtgccc	cacagcagca	ggtgttgagc	atgggctgag	120
aagctggacc	ggcaccaaag	ggctggcaga	aatgggcgcc	tggctgattc	ctaggcagtt	180
ggcggcagca	aggaggagag	gccgcagctt	ctggagcaga	gccgagacga	agcagttctg	240
gagtgcctga	acggccccct	gagccctacc	cgcctggccc	actatggtec	agaggctgtg	300
ggtgagccgc	ctgctgcggc	accggaaagc	ccagctcttg	ctgggtcaacc	tgctaacctt	360
tggcctggag	gtgtgttttg	ccgcaggcat	cacctatgtg	ccgcctctgc	tgctggaagt	420
gggggtagag	gagaagttca	tgaccatggt	gctgggcatt	ggtccagtgc	tgggcctggt	480
ctgtgtcccg	ctcttaggct	cagccagtga	ccactggcgt	ggacgctatg	gccgcgcccg	540
gcccttcac	tgggcaactgt	ccttgggcat	cctgctgagc	ctctttctca	tcccaagggc	600
cggctggcta	gcagggctgc	tgtgcccgga	tcccaggccc	ctggagctgg	cactgctcat	660
cctgggcgtg	gggtgctgtg	acttctgttg	ccagggtgtg	ttactccac	tggaggccct	720
gctctctgac	ctcttcgggg	accgcgacca	ctgtcgccag	gcctactctg	tctatgcctt	780
catgatcagt	cttgggggct	gcctgggcta	cctcctgcct	gccattgact	gggacaccag	840
tgccctggcc	ccctacctgg	gcacccagga	ggagtgcctc	tttggcctgc	tcacctcat	900
cttcctcacc	tgcgtagcag	ccacactgct	ggtggctgag	gaggcagcgc	tgggccccac	960
cgagccagca	gaagggtgtg	cggccccctc	cttgtcgccc	cactgctgtc	catgccgggc	1020
ccgcttggct	ttccggaacc	tgggcgcctt	gcttcccccg	ctgcaccagc	tgtgctgccg	1080
catgccccgc	accctgcgcc	ggctcttcgt	ggctgagctg	tgcaagctgga	tggcactcat	1140
gaccttcacg	ctgtttttaca	cggatttcgt	gggcgagggg	ctgtaccagg	gcgtgccccag	1200
agctgagccg	ggcaccgagg	cccggagaca	ctatgatgaa	ggcgttcgga	tgggcagcct	1260
ggggctgttc	ctgcagtgcg	ccatctccct	ggtcttctct	ctgggtcatg	accggctggt	1320
gcagcgattc	ggcactcgag	cagtctatatt	ggccagtgtg	gcagctttcc	ctgtggctgc	1380
cggtgccaca	tgctgtccc	acagtgtggc	cgtgggtgaca	gcttcagccg	ccctcaccgg	1440
gttcaccttc	tcagccctgc	agatcctgcc	ctacacactg	gcctccctct	accaccggga	1500
gaagcaggtg	ttcctgcccc	aataccgagg	ggacactgga	ggtgctagca	gtgaggacag	1560
cctgatgacc	agcttcctgc	caggccctaa	gcctggagct	cccttcctta	atggacacgt	1620
gggtgctgga	ggcagtggcc	tgtctcccac	tccaccgcg	ctctgcgggg	cctctgcctg	1680
tgatgtctcc	gtacgtgtgg	tgggtgggtga	gcccaccgag	gccaggggtg	ttccgggccg	1740
gggcatctgc	ctggacctcg	ccatcctgga	tagtgctctc	ctgctgtccc	aggtggcccc	1800
atccctgttt	atgggctcca	ttgtccagct	cagccagtct	gtcactgcct	atatggtgtc	1860
tgccgcaggc	ctgggtcttg	tgcctattta	ctttgtctaca	caggtagtat	ttgacaagag	1920



cgacttggcc	aaatactcag	cgtagaaaac	ttccagcaca	ttgggggtgga	gggcctgcct	1980
cactgggtcc	cagctccccg	ctcctgttag	ccccatgggg	ctgccgggct	ggccgccagt	2040
ttctgttgct	gccaaagtaa	tgtggctctc	tgtgccacc	ctgtgctgct	gaggtgcgta	2100
gctgcacagc	tgggggctgg	ggcgtccctc	tcctctctcc	ccagtctcta	gggctgcctg	2160
actggaggcc	ttccaagggg	gtttcagtc	ggacttatac	agggaggcca	gaagggtcc	2220
atgcactgga	atgcggggac	tctgcagggtg	gattaccag	gctcagggtt	aacagctagc	2280
ctcctagttg	agacacacct	agagaagggt	ttttgggagc	tgaataaact	cagtcacctg	2340
gtttcccatc	tctaagcccc	ttaacctgca	gcttcgttta	atgtagctct	tgcattgggag	2400
tttctaggat	gaaacactcc	tccatgggat	ttgaacatat	gacttatttg	taggggaaga	2460
gtcctgaggg	gcaacacaca	agaaccaggt	ccccccagcc	cacagcactg	tctttttgct	2520
gatccacccc	cctcttacct	tttatcagga	tgtggcctgt	tggtccttct	gttgccatca	2580
cagagacaca	ggcatttaaa	tatttaactt	atttatttaa	caaagtagaa	gggaatccat	2640
tgctagcttt	tctgtgttgg	tgtctaatat	ttgggtaggg	tgggggatcc	ccaacaatca	2700
gggtccctga	gatagctggt	cattgggctg	atcattgcca	gaatcttctt	ctcctgggggt	2760
ctggccccc	aaaatgccta	accaggacc	ttggaaattc	tactcatccc	aaatgataat	2820
tccaaatgct	gttacccaag	gttaggggtg	tgaaggaagg	tagaggggtg	ggcttcagggt	2880
ctcaacggct	tccctaacca	cccctcttct	cttggcccag	cctgggttccc	cccacttcca	2940
ctccctcta	ctctctctag	gactgggctg	atgaaggcac	tgcccaaaat	ttccctacc	3000
cccaactttc	ccctaccccc	aactttcccc	accagctcca	caacctgtt	tggagctact	3060
gcaggaccag	aagcacaaag	tgcggtttcc	caagcctttg	tccatctcag	ccccagagt	3120
atatctgtgc	ttggggaatc	tcacacagaa	actcaggagc	acccctgccc	tgagctaagg	3180
gaggtcttat	ctctcagggg	gggtttaagt	gcggtttgca	ataatgtcgt	cttattttatt	3240
tagcggggtg	aatattttat	actgtaagtg	agcaatcaga	gtataatgtt	tatgggtgaca	3300
aaattaaagg	ctttcttata	tgtttaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3360
aaaaaaaaara	aaaaaaaaaa	aaaaaaaaaa	aaaaaaataa	aaaaaaaaaa		3410

&lt;210&gt; 111

&lt;211&gt; 1289

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 111

agccaggcgt	ccctctgcct	gccactcag	tggcaacacc	cgaggagctgt	tttgtccttt	60
gtggagcctc	agcagttccc	tctttcagaa	ctcactgcca	agagccctga	acaggagcca	120
ccatgcagtg	cttcagcttc	attaagacca	tgatgatcct	cttcaatttg	ctcatctttc	180
tgtgtggtgc	agccctgttg	gcagtgggca	tctgggtgtc	aatcgatggg	gcaccccttc	240
tgaagatctt	cgggccactg	tcgtccagtg	ccatgcagtt	tgtcaacgtg	ggctacttcc	300
tcacgcagc	cggcggtgtg	gtctttgtct	ttggtttcc	gggctgctat	ggtgctaaga	360
ctgagagcaa	gtgtgccctc	gtgacgttct	tcttcaccc	cctccctcat	ttcattgctg	420
aggttgcagc	tgctgtggtc	gccttggtgt	acaccacaat	ggctgagcac	ttcctgacgt	480
tgctggtagt	gcctgccatc	aagaaagatt	atgggttcca	ggaagacttc	actcaagtgt	540
ggaacaccac	catgaaagg	ctcaagtgt	gtggcttcac	caactatacg	gattttgagg	600
actcacccta	cttcaaagag	aacagtgcct	ttccccatt	ctgttgcaat	gacaacgtca	660
ccaacacagc	caatgaaacc	tgcaccaagc	aaaaggctca	cgacaaaaaa	gtagagggtt	720
gcttcaatca	gcttttgtat	gacatccgaa	ctaattgcagt	caccgtgggt	ggtgtggcag	780
ctggaattgg	gggcctcgag	ctggctgcca	tgattgtgtc	catgtatctg	tactgcaatc	840
tacaataagt	ccacttctgc	ctctgccact	actgctgcca	catgggaact	gtgaagaggc	900
accctggcaa	gcagcagtg	ttgggggagg	ggacaggatc	taacaatgtc	acttgggcca	960
gaatggacct	gccctttctg	ctccagactt	ggggctagat	agggaccact	ccttttagcg	1020
atgctgact	ttccttccat	tggtgggtgg	atgggtgggg	ggcattccag	agcctctaag	1080
gtagccagtt	ctgttgccca	ttccccag	ctattaaacc	cttgatatgc	cccctaggcc	1140



```

tagtgggtgat cccagtgctc tactggggga tgagagaaag gcattttata gcctgggcat 1200
aagtgaaatc agcagagcct ctgggtggat gtgtagaagg cacttcaaaa tgcataaacc 1260
tgttacaatg ttaaaaaaaaa aaaaaaaaaa 1289

```

```

<210> 112
<211> 315
<212> PRT
<213> Homo sapien

```

```

<400> 112
Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val Asn Lys Gln
 1          5          10          15
Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys Asp Val Phe
 20          25          30
Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr Gly Val Ala
 35          40          45
Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro Ser Ile Leu
 50          55          60
Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro
 65          70          75          80
Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn Cys Ser Ser
 85          90          95
Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys
100          105          110
Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe
115          120          125
Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe
130          135          140
Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys
145          150          155          160
Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala Leu
165          170          175
Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu Arg Gln
180          185          190
Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro Ala Leu Glu
195          200          205
His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys Leu Leu Thr
210          215          220
Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg Ala Arg Asp
225          230          235          240
Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys Arg Thr Ser Gln Lys Val
245          250          255
Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr Glu Gln Arg
260          265          270
Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Ser Arg Val Leu Gly
275          280          285
Trp Val Ala Glu Ala Leu Ser Arg Ser Ala Leu Leu Pro Pro Gly Gly
290          295          300
Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp
305          310          315

```

CC-0.2019-01-10



<400> 113															
Met 1	Val	Gln	Arg	Leu 5	Trp	Val	Ser	Arg	Leu 10	Leu	Arg	His	Arg	Lys 15	Ala
Gln	Leu	Leu	Leu	Val 20	Asn	Leu	Leu	Thr	Phe	Gly	Leu	Glu	Val	Cys	Leu
Ala	Ala	Gly	Ile	Thr 35	Tyr	Val	Pro	Pro	Leu	Leu	Leu	Glu	Val	Gly	Val
Glu	Glu	Lys	Phe	Met 50	Thr	Met	Val	Leu	Gly	Ile	Gly	Pro	Val	Leu	Gly
Leu 65	Val	Cys	Val	Pro 70	Leu	Leu	Gly	Ser	Ala	Ser	Asp	His	Trp	Arg	Gly
Arg	Tyr	Gly	Arg	Arg 85	Arg	Pro	Phe	Ile	Trp	Ala	Leu	Ser	Leu	Gly	Ile
Leu	Leu	Ser	Leu	Phe 100	Leu	Ile	Pro	Arg	Ala	Gly	Trp	Leu	Ala	Gly	Leu
Leu	Cys	Pro	Asp	Pro 115	Arg	Pro	Leu	Glu	Leu	Ala	Leu	Leu	Ile	Leu	Gly
Val	Gly	Leu	Leu	Asp 130	Phe	Cys	Gly	Gln	Val	Cys	Phe	Thr	Pro	Leu	Glu
Ala 145	Leu	Leu	Ser	Asp 150	Leu	Phe	Arg	Asp	Pro	Asp	His	Cys	Arg	Gln	Ala
Tyr	Ser	Val	Tyr	Ala 165	Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu	Gly	Tyr
Leu	Leu	Pro	Ala	Ile 180	Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro	Tyr	Leu
Gly	Thr	Gln	Glu	Glu 195	Cys	Leu	Phe	Gly	Leu	Leu	Thr	Leu	Ile	Phe	Leu
Thr	Cys	Val	Ala	Ala 210	Thr	Leu	Leu	Val	Ala	Glu	Glu	Ala	Ala	Leu	Gly
Pro 225	Thr	Glu	Pro	Ala 230	Glu	Gly	Leu	Ser	Ala	Pro	Ser	Leu	Ser	Pro	His
Cys	Cys	Pro	Cys	Arg 245	Ala	Arg	Leu	Ala	Phe	Arg	Asn	Leu	Gly	Ala	Leu
Leu	Pro	Arg	Leu	His 260	Gln	Leu	Cys	Cys	Arg	Met	Pro	Arg	Thr	Leu	Arg
Arg	Leu	Phe	Val	Ala 275	Glu	Leu	Cys	Ser	Trp	Met	Ala	Leu	Met	Thr	Phe
Thr	Leu	Phe	Tyr	Thr 290	Asp	Phe	Val	Gly	Glu	Gly	Leu	Tyr	Gln	Gly	Val
Pro 305	Arg	Ala	Glu	Pro 310	Gly	Thr	Glu	Ala	Arg	Arg	His	Tyr	Asp	Glu	Gly
Val	Arg	Met	Gly	Ser 325	Leu	Gly	Leu	Phe	Leu	Gln	Cys	Ala	Ile	Ser	Leu
Val	Phe	Ser	Leu	Val 340	Met	Asp	Arg	Leu	Val	Gln	Arg	Phe	Gly	Thr	Arg
Ala	Val	Tyr	Leu	Ala 345	Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala	Gly	Ala







145		150		155		160									
Ser	Pro	Tyr	Phe	Lys	Glu	Asn	Ser	Ala	Phe	Pro	Pro	Phe	Cys	Cys	Asn
				165					170					175	
Asp	Asn	Val	Thr	Asn	Thr	Ala	Asn	Glu	Thr	Cys	Thr	Lys	Gln	Lys	Ala
			180					185					190		
His	Asp	Gln	Lys	Val	Glu	Gly	Cys	Phe	Asn	Gln	Leu	Leu	Tyr	Asp	Ile
		195					200					205			
Arg	Thr	Asn	Ala	Val	Thr	Val	Gly	Gly	Val	Ala	Ala	Gly	Ile	Gly	Gly
	210					215				220					
Leu	Glu	Leu	Ala	Ala	Met	Ile	Val	Ser	Met	Tyr	Leu	Tyr	Cys	Asn	Leu
225					230				235					240	
Gln															

<210> 115  
 <211> 366  
 <212> DNA  
 <213> Homo sapien

<400> 115	
gctctttctc tcccctctc tgaatttaat tctttcaact tgcaatttgc aaggattaca	60
catttcactg tgatgtatat tgtgttgcaa aaaaaaaaaa gtgtctttgt ttaaaattac	120
ttggtttggt aatccatctt gctttttccc catttgaact agtcattaac ccatctctga	180
actggtagaa aaacatctga agagctagtc tatcagcatc tgacaggtga attggatggt	240
tctcagaacc atttcaccca gacagcctgt ttctatcctg ttttaataaat tagtttgggt	300
tctctacatg cataacaaac cctgctccaa tctgtcacat aaaagtctgt gacttgaagt	360
ttagtc	366

<210> 116  
 <211> 282  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(282)  
 <223> n = A,T,C or G

<400> 116	
acaaagatga accattttcct atattatagc aaaattaaaa tctaccgta ttctaattatt	60
gagaaatgag atnaaacaca atnttataaa gtctacttag agaagatcaa gtgacctcaa	120
agactttact attttcatat ttttaagacac atgattttatc ctatttttagt aacctgggtc	180
atacgttaaa caaaggataa tgtgaacagc agagaggatt tgttggcaga aaatctatgt	240
tcaatctnga actatctana tcacagacat ttctatttct tt	282

<210> 117  
 <211> 305  
 <212> DNA  
 <213> Homo sapien

<220>



<221> misc\_feature  
 <222> (1)...(305)  
 <223> n = A,T,C or G

<400> 117  
 acacatgtcg cttcactgcc ttcttagatg cttctgggtca acatanagga acagggacca 60  
 tatttattcct ccctcctgaa acaattgcaa aataanacaa aatatatgaa acaattgcaa 120  
 aataaggcaa aatatatgaa acaacagggtc tcgagatatt ggaaatcagt caatgaagga 180  
 tactgatccc tgatcactgt cctaattgcag gatgtgggaa acagatgagg tcacctctgt 240  
 gactgccccca gcttactgcc tgtagagagt ttctangctg cagttcagac agggagaaat 300  
 tgggt 305

<210> 118  
 <211> 71  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(71)  
 <223> n = A,T,C or G

<400> 118  
 accaaggtgt ntgaatctct gacgtgggga tctctgattc ccgcacaatc tgagtggaaa 60  
 aantcctggg t 71

<210> 119  
 <211> 212  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(212)  
 <223> n = A,T,C or G

<400> 119  
 actccggttg gtgtcagcag cacgtggcat tgaacatngc aatgtggagc ccaaaccaca 60  
 gaaaatgggg tgaaattggc caactttcta tnaacttatg ttggcaantt tgccaccaac 120  
 agtaagctgg cccttctaataaaaagaaaat tgaaaggttt ctcactaanc ggaattaant 180  
 aatggantca aganactccc aggcctcagc gt 212

<210> 120  
 <211> 90  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(90)  
 <223> n = A,T,C or G



<400> 120  
 actcgttgca natcaggggc cccccagagt caccgttgca ggagtccttc tggctcttgcc 60  
 ctccgccggc gcagaacatg ctgggggtggt 90

<210> 121  
 <211> 218  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(218)  
 <223> n = A,T,C or G

<400> 121  
 tgtancgtga anacgacaga naggggtgtc aaaaatggag aanccttgaa gtcattttga 60  
 gaataagatt tgctaaaaga tttggggcta aaacatgggtt attgggagac atttctgaag 120  
 atatncangt aaattangga atgaattcat gggttcttttg ggaattcctt tacgatngcc 180  
 agcatanact tcatgtgggg atancagcta cccttgta 218

<210> 122  
 <211> 171  
 <212> DNA  
 <213> Homo sapien

<400> 122  
 taggggtgta tgcaactgta aggacaaaaa ttgagactca actggcttaa ccaataaagg 60  
 ctttgttag ctcatggaac aggaagtcgg atgggtggggc atcttcagtg ctgcatgagt 120  
 caccaccccg gcgggggtcat ctgtgccaca ggtccctggt gacagtgcgg t 171

<210> 123  
 <211> 76  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(76)  
 <223> n = A,T,C or G

<400> 123  
 tgtagcgtga agacnacaga atgggtgtgtg ctgtgctatc caggaacaca tttattatca 60  
 ttatcaanta ttgtgt 76

<210> 124  
 <211> 131  
 <212> DNA  
 <213> Homo sapien

<400> 124



```

acctttcccc aaggccaatg tctgtgtgc taactggccg gctgcaggac agctgcaatt    60
caatgtgctg ggtcatatgg aggggaggag actctaaaat agccaatttt attctcttgg    120
ttaagatttg t                                     131

```

```

<210> 125
<211> 432
<212> DNA
<213> Homo sapien

```

```

<400> 125
actttatcta ctggctatga aatagatggg ggaaaattgc gttaccaact ataccactgg    60
cttgaaaaag aggtgatagc tcttcagagg acttgtgact tttgctcaga tgctgaagaa    120
ctacagtctg catttggcag aaatgaagat gaatttggat taaatgagga tgctgaagat    180
ttgcctcacc aaacaaaagt gaaacaactg agagaaaatt ttcaggaaaa aagacagtgg    240
ctcttgaagt atcagtcact tttgagaatg tttcttagtt actgcatact tcatggatcc    300
catggtgggg gtcttgcatc tgtaagaatg gaattgattt tgcttttgca agaatctcag    360
caggaaacat cagaaccact attttctagc cctctgtcag agcaaaccctc agtgcctctc    420
ctctttgctt gt                                     432

```

```

<210> 126
<211> 112
<212> DNA
<213> Homo sapien

```

```

<400> 126
acacaacttg aatagtaaaa tagaaactga gctgaaattt ctaattcact ttctaaccat    60
agtaagaatg atatttcccc ccagggatca ccaaataatt ataaaaattt gt          112

```

```

<210> 127
<211> 54
<212> DNA
<213> Homo sapien

```

```

<400> 127
accacgaaac cacaacaag atggaagcat caatccactt gccaaagcaca gcag          54

```

```

<210> 128
<211> 323
<212> DNA
<213> Homo sapien

```

```

<400> 128
acctcattag taattgtttt gttgtttcat ttttttctaa tgtctcccct ctaccagctc    60
acctgagata acagaatgaa aatggaagga cagccagatt tctcctttgc tctctgctca    120
ttctctctga agtctaggtt acccattttg gggaccatt ataggcaata aacacagttc    180
ccaaagcatt tggacagttt cttgttgtgt tttagaatgg ttttcctttt tcttagcctt    240
ttcctgcaa aggtcactc agtcccttgc ttgctcagtg gactgggctc cccagggcct    300
aggctgcctt cttttccatg tcc                                     323

```

```

<210> 129
<211> 192

```



<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(192)  
<223> n = A,T,C or G

<400> 129  
acatacatgt gtgtatatatt ttaaatatca cttttgtatc actctgactt tttagcatatc 60  
tgaaaacaca ctaacataat ttntgtgaac catgatcaga tacaacccaa atcattcatc 120  
tagcacattc atctgtgata naaagatagg tgagtttcat ttccttcacg ttggccaatg 180  
gataaacaaa gt 192

<210> 130  
<211> 362  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(362)  
<223> n = A,T,C or G

<400> 130  
ccctttttta tggaatgagt agactgtatg tttgaanatt tanccacaac ctctttgaca 60  
tataatgacg caacaaaaag gtgctgttta gtcctatggg tcagtttatg ccctgacaa 120  
gtttccattg tgttttgccg atcttctggc taatcgtggg atcctccatg ttattagtaa 180  
ttctgtattc ctttttggtta acgcctggta gatgtaacct gctangaggc taactttata 240  
cttattttaa agctcttatt ttgtgggtcat taaaatggca atttatgtgc agcactttat 300  
tgcagcagga agcacgtgtg gggttggttg aaagctcttt gctaattctta aaaagtaatg 360  
gg 362

<210> 131  
<211> 332  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(332)  
<223> n = A,T,C or G

<400> 131  
ctttttgaaa gatcgtgtcc actcctgtgg acatcttggt ttaatggagt ttcccatgca 60  
gtangactgg tatggttgca gctgtccaga taaaaacatt tgaagagctc caaatgaga 120  
gttctcccag gttcgccctg ctgctccaag tctcagcagc agcctctttt aggaggcatc 180  
ttctgaacta gattaaggca gcttgtaaat ctgatgtgat ttggtttatt atccaactaa 240  
cttccatctg ttatcactgg agaaagccca gactcccan gacnggtacg gattgtgggc 300  
atanaaggat tgggtgaagc tggcgttgtg gt 332



<210> 135



<211> 350  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(350)  
 <223> n = A,T,C or G

<400> 135  
 acttanaacc atgcctagca catcagaatc cctcaaagaa catcagtata atcctataacc 60  
 atancaagtg gtgactgggt aagcgtgcga caaagggtcag ctggcacatt acttgtgtgc 120  
 aaacttgata cttttgttct aagtaggaac tagtatacag tncctaggan tgggtactcca 180  
 ggggtgcccc caactcctgc agcgcctcct ctgtgccagn ccctgnaagg aactttcgtc 240  
 ccacctcaat caagccctgg gccatgctac ctgcaattgg ctgaacaaac gtttgctgag 300  
 ttccaagga tgcaagcct ggtgctcaac tcttggggcg tcaactcagt 350

<210> 136  
 <211> 399  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(399)  
 <223> n = A,T,C or G

<400> 136  
 tgtacctga agacgacaga agttgcatgg cagggacagg gcagggccga ggccaggggt 60  
 gctgtgattg tatccgaata ntcctcgtga gaaaagataa tgagatgacg tgagcagcct 120  
 gcagacttgt gtctgccttc aanaagccag acaggaaggc cctgcctgcc ttggctctga 180  
 cctggcgccc agccagccag ccacaggtgg gcttcttcct tttgtggtga caacnccaag 240  
 aaaactgcag agggcccagg tccaggtgtna gtgggtangt gaccataaaa caccaggtgc 300  
 tcccaggaac cggggcaaag gccatcccca cctacagcca gcatgcccac tggcgtgatg 360  
 ggtgcagang gatgaagcag ccagntgttc tgctgtggt 399

<210> 137  
 <211> 165  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(165)  
 <223> n = A,T,C or G

<400> 137  
 actggtgtgg tnggggggtga tgctgggtgg anaagttgan gtgacttcan gatggtgtgt 60  
 ggaggaagtg tgtgaacgta gggatgtaga ngttttggcc gtgctaaatg agcttcggga 120  
 ttggtcgggt ccactgggtg tcaactgtcat tgggtggggt cctgt 165



<210> 138  
 <211> 338  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(338)  
 <223> n = A,T,C or G

<400> 138  
 actcactgga atgccacatt cacaacagaa tcagaggtct gtgaaaacat taatggctcc 60  
 ttaacttctc cagtaagaat cagggacttg aaatggaaac gttaacagcc acatgcccaa 120  
 tgctgggcag tctcccatgc cttccacagt gaaagggctt gagaaaaatc acatccaatg 180  
 tcatgtgttt ccagccacac caaaaggtgc ttgggggtgga gggctggggg catananggt 240  
 cangcctcag gaagcctcaa gttccattca gctttgccac tgtacattcc ccatntttaa 300  
 aaaaactgat gccttttttt ttttttttg taaaattc 338

<210> 139  
 <211> 382  
 <212> DNA  
 <213> Homo sapien

<400> 139  
 gggaatcttg gtttttggca tctggtttgc ctatagccga ggccactttg acagaacaaa 60  
 gaaagggact tcgagtaaga aggtgattta cagccagcct agtgcccga gtgaaggaga 120  
 attcaaacag acctcgatcat tcttggtgtg agcctggctg gctcaccgcc tatcatctgc 180  
 atttgcttta ctcaggtgct accggactct ggcccctgat gtctgtagtt tcacaggatg 240  
 ccttattttgt cttctacacc ccacagggcc ccctacttct tcggatgtgt ttttaataat 300  
 gtcagctatg tgcccatcc tcttcatgc cctccctccc tttcctacca ctgctgagtg 360  
 gcctggaact tgtttaaagt gt 382

<210> 140  
 <211> 200  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(200)  
 <223> n = A,T,C or G

<400> 140  
 accaaaanctt ctttctgttg tggtngattt tactataggg gtttngcttn ttctaaanat 60  
 acttttcatt taacancttt tggttaagtgt caggctgcac tttgctccat anaattattg 120  
 ttttcacatt tcaacttgta tgtgtttgtc tcttanagca ttggtgaaat cacatatttt 180  
 atattcagca taaaggagaa 200

<210> 141  
 <211> 335  
 <212> DNA



<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(335)

<223> n = A,T,C or G

<400> 141

```
actttatttt caaaacactc atatgttgca aaaaacacat agaaaaataa agtttggtgg      60
gggtgctgac taaacttcaa gtcacagact tttatgtgac agattggagc agggtttggt      120
atgcatgtag agaaccctaa ctaatttatt aaacaggata gaaacaggct gtctgggtga      180
aatggttctg agaaccatcc aattcacctg tcagatgctg atanactagc tcttcagatg      240
tttttctacc agttcagaga tnggttaatg actanttcca atgggggaaaa agcaagatgg      300
attcacaac caagtaattt taaacaaaga cactt                                     335
```

<210> 142

<211> 459

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(459)

<223> n = A,T,C or G

<400> 142

```
accaggttaa tattgccaca tatatccttt ccaattgctg gctaaacaga cgtgtattta      60
gggttggtta aagacaaccc agcttaatat caagagaaat tgtgaccttt catggagtat      120
ctgatggaga aaacactgag ttttgacaaa tcttatttta ttcagatagc agtctgatca      180
cacatgggcc aacaacactc aaataataaa tcaaataatna tcagatgtta aagattggct      240
ttcaaacatc atagccaatg atgccccgct tgcctataat ctctccgaca taaaaccaca      300
tcaacacctc agtggccacc aaaccattca gcacagcttc cttaactgtg agctgtttga      360
agctaccagt ctgagcacta ttgactatnt ttttcangct ctgaatagct ctagggatct      420
cagcangggg gggaggaacc agctcaacct tggcgtant                                     459
```

<210> 143

<211> 140

<212> DNA

<213> Homo sapien

<400> 143

```
acatttcctt ccaccaagtc aggactcctg gcttctgtgg gagttcttat cacctgaggg      60
aaatccaaac agtctctcct agaaaggaat agtgtcacca accccaccca tctccctgag      120
accatccgac ttccctgtgt                                     140
```

<210> 144

<211> 164

<212> DNA

<213> Homo sapien

<220>



<221> misc\_feature  
 <222> (1)...(164)  
 <223> n = A,T,C or G

<400> 144  
 acttcagtaa caacatacaa taacaacatt aagtgtatat tgccatcttt gtcattttct 60  
 atctatacca ctctcccttc tgaaaacaan aatcactanc caatcactta taaaaatttg 120  
 aggcaattaa tccatatttg ttttcaataa ggaaaaaaag atgt 164

<210> 145  
 <211> 303  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(303)  
 <223> n = A,T,C or G

<400> 145  
 acgtagacca tccaactttg tatttgtaat ggcaaacatc cagnagcaat tcctaaacaa 60  
 actggagggt atttataccc aattatccca ttcattaaca tgccctcctc ctcaggctat 120  
 gcaggacagc tatcataagt cggcccaggc atccagatac taccatttgt ataaacttca 180  
 gtaggggagt ccatccaagt gacaggtcta atcaaaggag gaaatggaac ataagcccag 240  
 tagtaaaatn ttgcttagct gaaacagcca caaaagactt accgccgtgg tgattaccat 300  
 caa 303

<210> 146  
 <211> 327  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(327)  
 <223> n = A,T,C or G

<400> 146  
 actgcagctc aattagaagt ggtctctgac tttcatcanc ttctccctgg gctccatgac 60  
 actggcctgg agtgactcat tgctctgggt ggttgagaga gtccttttgc caacaggcct 120  
 ccaagtcagg gctgggattt gtttcctttc cacattctag caacaatatg ctggccactt 180  
 cctgaacagg gaggggtgga ggagccagca tggaacaagc tgccactttc taaagtagcc 240  
 agacttgccc ctgggcctgt cacacctact gatgaccttc tgtgcctgca ggatggaatg 300  
 taggggtgag ctgtgtgact ctatgggt 327

<210> 147  
 <211> 173  
 <212> DNA  
 <213> Homo sapien

<220>



<221> misc\_feature  
 <222> (1)...(173)  
 <223> n = A,T,C or G

<400> 147  
 acattgtttt tttgagataa agcattgana gagctctcct taacgtgaca caatggaagg 60  
 actggaacac ataccacat ctttgttctg agggataatt ttctgataaa gtcttgctgt 120  
 atattcaagc acatatgtta tatattattc agttccatgt ttatagccta gtt 173

<210> 148  
 <211> 477  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(477)  
 <223> n = A,T,C or G

<400> 148  
 acaaccactt tatctcatcg aatttttaac ccaaactcac tcaactgtgcc tttctatcct 60  
 atgggatata ttatttgatg ctccatttca tcacacatat atgaataata cactcatact 120  
 gccctactac ctgctgcaat aatcacattc ccttctctgtc ctgaccctga agccattggg 180  
 gtggtcctag tggccatcag tccangcctg caccttgagc ccttgagctc cattgctcac 240  
 nccanccac ctcaccgacc ccatactctt acacagctac ctccttgctc tctaacccca 300  
 tagattatnt ccaaattcag tcaattaagt tactattaac actctaccg acatgtccag 360  
 caccactggt aagccttctc cagccaacac acacacacac acacncacac acacacatat 420  
 ccaggcacag gctacctcat cttcacaatc acccctttaa ttacctgct atggtgg 477

<210> 149  
 <211> 207  
 <212> DNA  
 <213> Homo sapien

<400> 149  
 acagttgtat tataatatca agaaataaac ttgcaatgag agcatttaag agggaagaac 60  
 taacgtatatt tagagagcca aggaaggttt ctgtgggggag tgggatgtaa ggtggggcct 120  
 gatgataaat aagagtcagc caggtaagtg ggtggtgtgg tatgggcaca gtgaagaaca 180  
 tttcaggcag agggaacagc agtgaaa 207

<210> 150  
 <211> 111  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(111)  
 <223> n = A,T,C or G

<400> 150



accttgattt cattgctgct ctgatggaaa cccaactatc taatttagct aaaacatggg 60  
cacttaaagt tggtcagtgt ttggacttgt taactantgg catctttggg t 111

<210> 151  
<211> 196  
<212> DNA  
<213> Homo sapien

<400> 151  
agcgcggcag gtcattattga acattccaga tacctatcat tactcgatgc tgttgataac 60  
agcaagatgg ctttgaactc agggtcacca ccagctattg gaccttacta tgaaaaccat 120  
ggataccaac cggaaaaccc ctatcccgcg cagcccactg tgggtcccccac tgtctacgag 180  
gtgcatccgg ctccagt 196

<210> 152  
<211> 132  
<212> DNA  
<213> Homo sapien

<400> 152  
acagcacttt cacatgtaag aagggagaaa ttcctaaatg taggagaaag ataacagAAC 60  
cttccccttt tcatctagtg gtggaaacct gatgctttat gttgacagga atagaaccag 120  
gagggagtgt gt 132

<210> 153  
<211> 285  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1) ... (285)  
<223> n = A,T,C or G

<400> 153  
acaanaccca nganaggcca ctggccgtgg tgtcatggcc tccaaacatg aaagtgtcag 60  
cttctgctct tatgtcctca tctgacaact ctttaccatt tttatcctcg ctcagcagga 120  
gcacatcaat aaagtccaaa gtcttggact tggccttggc ttggaggaag tcatcaacac 180  
cctggctagt gaggggtgagg cgccgctcct ggatgacggc atctgtgaag tcgtgcacca 240  
gtctgcaggc cctgtggaag cgccgtccac acggagtnag gaatt 285

<210> 154  
<211> 333  
<212> DNA  
<213> Homo sapien

<400> 154  
accacagtcc tgttgggcca gggettcatt accctttctg tgaaaagcca tattatcacc 60  
accccaaatt tttccttaaa tatctttaac tgaaggggtc agcctcttga ctgcaaagac 120  
cctaagccgg ttacacagct aactcccact ggccctgatt tgtgaaattg ctgctgcctg 180  
attggcacag gagtcgaagg tgttcagctc cctcctccg tggaacgaga ctctgatttg 240



```
<210> 155
<211> 308
<212> DNA
<213> Homo sapien
```

<400> 155

```
<210> 156
<211> 295
<212> DNA
<213> Homo sapien
```

<400> 156

```
<210> 157
<211> 126
<212> DNA
<213> Homo sapien
```

<400> 157

```
<210> 158
<211> 442
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc_feature  
<222> (1)...(442)
```



acctgcaccc	agcttccctg	ccaaactcac	aaggagacat	caacctctag	acagggaaac	60
agcttcagga	tacttccagg	agacagagcc	accagcagca	aaacaaatat	tcccatgcct	120
ggagcatggc	atagaggaag	ctganaaatg	tggggtctga	ggaagccatt	tgagtctggc	180
cactagacat	ctcatcagcc	acttgtgtga	agagatgcc	catgacccca	gatgcctctc	240
ccacccttac	ctccatctca	cacacttgag	ctttccactc	tgtataattc	taacatcctg	300
gagaaaaaatg	gcagtttgac	cgaacctgtt	cacaacggta	gaggctgatt	tctaacgaaa	360
cttgtagaat	qaagcctgga					380



<400> 161

<210> 162

<400> 162

<210> 163

<220>

<400> 163

<210> 164

<220>

<400> 164

cttatcacaa	tgaatgttct	cctgggcagc	gttgatgatct	ttgccacctt	cgtgacttta	60
tgcaatgcat	catgctattt	catacctaatt	gagggagtct	caggagattc	aaccaggaaa	120
tgcatggatc	tcaaaggaaa	caaacaccca	ataaactcgg	agtggcagac	tgacaactgt	180
gagacatgca	cttgctacga	aacagaaatt	tcattgttgca	cccttgtttc	tacacctgtg	240
ggttatgaca	aagacaactg	ccaaagaatc	ttcaagaagg	aggactgcaa	gtatatcgtg	300



```
<210> 165
<211> 195
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(195)
<223> n = A,T,C or G
```

```
<210> 166
<211> 383
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(383)
<223> n = A,T,C or G
```

```
<210> 167
<211> 247
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(247)
<223> n = A,T,C or G
```

<400> 167  
acagagccag accttgGCCA taaatgaanc agagattaag actaaacccc aagtcganat 60



```
<210> 168
<211> 273
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(273)
<223> n = A,T,C or G
```

```
<210> 169
<211> 431
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(431)
<223> n = A,T,C or G
```

```
<210> 170
<211> 266
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(266)
<223> n = A,T,C or G
```



&lt;400&gt; 170

```

acctgtgggc tgggctgtta tgccctgtgcc ggctgctgaa agggagttca gaggtggagc      60
tcaaggagct ctgcaggcat tttgccaanc ctctccanag canagggagc aacctacact      120
ccccgctaga aagacaccag attggagtcc tgggaggggg agttgggggtg ggcatttgat      180
gtatacttgt cacctgaatg aangagccag agaggaanga gacgaanatg anattggcct      240
tcaaagctag gggctctggca ggtgga                                     266

```

&lt;210&gt; 171

&lt;211&gt; 1248

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(1248)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 171

```

ggcagccaaa tcataaacgg cgaggactgc agcccgcaact cgcagccctg gcaggcggca      60
ctggtcatgg aaaacgaatt gttctgctcg ggcgtcctgg tgcattccga gtgggtgctg      120
tcagccgcac actgtttcca gaagtgaagt cagagctcct acaccatcgg gctgggcctg      180
cacagtcttg aggccgacca agagccaggg agccagatgg tggaggccag cctctccgta      240
cggcaccag agtacaacag acccttgctc gctaacgacc tcatgctcat caagttggac      300
gaatccgtgt ccgagtctga caccatccgg agcatcagca ttgcttcgca gtgccctacc      360
gcggggaact cttgcctcgt ttctggctgg ggtctgctgg cgaacggcag aatgcctacc      420
gtgctgcagt gcgtgaacgt gtcggtggtg tctgaggagg tctgcagtaa gctctatgac      480
ccgctgtacc accccagcat gttctgcgcc ggcggagggc aagaccagaa ggactcctgc      540
aacggtgact ctgggggggcc cctgatctgc aacgggtact tgcagggcct tgtgtctttc      600
ggaaaagccc cgtgtggcca agttggcgtg ccagggtgtct acaccaacct ctgcaaattc      660
actgagtgga tagagaaaac cgtccaggcc agttaactct ggggactggg aacccatgaa      720
attgaccccc aaatacatcc tgcggaagga attcaggaat atctgttccc agccccctcct      780
ccctcaggcc caggagtcca ggccccccagc ccctcctccc tcaaaccaag ggtacagatc      840
cccagccctt cctccctcag acccaggagt ccagaccccc cagccccctc tccctcagac      900
ccaggagtcc agccccctcct ccctcagacc caggagtcca gacccccccag cccctcctcc      960
ctcagaccca ggggtccagg ccccccaacc ctcctccctc agactcagag gtccaagccc     1020
ccaaccntc attccccaga ccagagggtc cagggtcccag cccctcntcc ctcagaccca     1080
gcggtccaat gccacctaga ctntccctgt acacagtgcc cccttgtggc acgttgaccc     1140
aaccttacca gttggttttt catttttngt ccctttcccc tagatccaga aataaagttt     1200
aagagaagng caaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaa                    1248

```

&lt;210&gt; 172

&lt;211&gt; 159

&lt;212&gt; PRT

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (1)...(159)

&lt;223&gt; Xaa = Any Amino Acid



ggcagccgc	actcgagcc	ctggcaggcg	gcactggtca	tggaaaacga	attgtttctgc	60
tcgggcgtcc	tggtgcatacc	gcagtggggtg	ctgtcagccg	cacactgttt	ccagaactcc	120
tacaccatcg	ggctgggect	gcacagtctt	gaggccgacc	aagagccagg	gagccagatg	180
gtggaggcca	gcctctccgt	acggcaccca	gagtacaaca	gaccttctgt	cgctaacgac	240
ctcatgtctca	tcaagttgga	cgaatccgtg	tccgagtctg	acaccatccg	gagcatcagc	300
attgcttcgc	agtgcctac	cgcggggaac	tcttgctctg	tttctggctg	gggtctgctg	360
gcgaacggtg	agctcacggg	tgtgtgtctg	ccctcttcaa	ggaggtcctc	tgcccagtcg	420
cgggggctga	cccagagctc	tgcgtcccag	gcagaatgcc	taccgtgctg	cagtgcgtga	480
acgtgtcggg	ggtgtctgag	gagggtctgca	gtaagctcta	tgacccgctg	taccacccca	540
gcatgttctg	cgccggcgga	gggcaagacc	agaaggactc	ctgcaacggg	gactctgggg	600
ggccctgat	ctgcaacggg	tacttgcaag	gccttgtgtc	tttcggaaaa	gccccgtgtg	660
gccaaagtgg	cgtgccagg	gtctacacca	acctctgcaa	attcactgag	tggaatagaga	720
aaaccgtcca	ggccagttaa	ctctggggac	tgggaaccca	tgaaattgac	cccaaaatac	780
atcctgcgga	aggaattcag	gaatatctgt	tcccagcccc	tctcctctca	ggcccaggag	840
tccaggcccc	cagccctcc	tccctcaaac	caagggtaca	gatccccagc	ccctcctccc	900
tcagaccag	gagtcagac	ccccagccc	ctctcctc	agaccagga	gtccagcccc	960
tctcctca	gaccagga	tccagcccc	ccagccctc	ctcctcaga	cccaggggtt	1020
gaggcccca	acccctctc	cttcagagtc	agagggtcaa	gcccccaacc	cctcgttccc	1080



```

cagacccaga ggtnnaggtc ccagcccctc ttcctcaga cccagnggtc caatgccacc 1140
tagatthttcc ctgnacacag tgcccccttg tggngngttg acccaacctt accagttggt 1200
ttttcatttt tngtcccttt cccctagatc cagaaataaa gtttaagaga ngngcaaaaa 1260
aaaaa 1265

```

```

<210> 174
<211> 1459
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(1459)
<223> n = A,T,C or G

```

```

<400> 174
ggtcagccgc acactgtttc cagaagtgag tgcagagctc ctacaccatc gggctggggc 60
tgcacagtct tgagggccgac caagagccag ggagccagat ggtggaggcc agcctctccg 120
tacggcacc agagtacaac agacccttgc tcgctaacga cctcatgctc atcaagttgg 180
acgaatccgt gtccgagtct gacaccatcc ggagcatcag cattgcttcg cagtgcctta 240
ccgcggggaa ctcttgctc gtttctggct ggggtctgct ggcgaacggt gagctcacgg 300
gtgtgtgtct gccctcttca aggaggtcct ctgccagtc gcgggggctg acccagagct 360
ctgctccca ggcagaatgc ctaccgtgct gcagtgcgtg aacgtgtcgg tgggtgtctga 420
ngaggtctgc antaagctct atgaccgct gtaccacccc ancatgttct gcgccggcgg 480
agggcaagac cagaaggact cctgcaacgt gagagagggg aaaggggagg gcaggcgact 540
cagggaaagg tggagaaggg ggagacagag acacacaggg ccgcatggcg agatgcagag 600
atggagagac acacagggag acagtgacaa ctagagagag aaactgagag aaacagagaa 660
ataaacacag gaataaagag aagcaaagga agagagaaac agaaacagac atggggaggc 720
agaaacacac acacatagaa atgcagttga ccttccaaca gcatggggcc tgagggcggt 780
gacctccacc caatagaaaa tctctttata acttttgact ccccaaaaac ctgactagaa 840
atagcctact gttgacgggg agccttacca ataacataaa tagtcgattt atgcatacgt 900
tttatgcatt catgatatac ctttggttga attttttgat atttctaagc tacacagttc 960
gtctgtgaat ttttttaaat tggtgcaact ctccataaat ttttctgatg tgtttattga 1020
aaaaatccaa gtataagtgg acttggtgcat tcaaaccagg gttgttcaag ggtcaactgt 1080
gtaccagag ggaaacagtg acacagattc atagaggtga aacacgaaga gaaacaggaa 1140
aatcaagac tctacaaaga ggctgggcag ggtggctcat gcctgtaatc ccagcacttt 1200
gggagggcag gcaggcagat cacttgaggt aaggagttca agaccagcct ggccaaaatg 1260
gtgaaatcct gtctgtacta aaaatacaaa agttagctgg atatggtggc aggcgcctgt 1320
aatccagct acttgggagg ctgaggcagg agaattgctt gaatatggga ggagaggtt 1380
gaagtgagtt gagatcacac cactatactc cagctggggc aacagagtaa gactctgtct 1440
caaaaaaaaa aaaaaaaaaa 1459

```

```

<210> 175
<211> 1167
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(1167)
<223> n = A,T,C or G

```



gogcagccct	ggcaggcggc	actggtcatg	gaaaacgaat	tgtttctgtc	gggcgtcctg	60
gtgcatccgc	agtgggtgct	gtcagccgca	cactgtttcc	agaactccta	caccatcggg	120
ctgggcctgc	acagtcttga	ggccgaccaa	gagccagggg	gccagatggg	ggaggccagc	180
ctctccgtac	ggcaccacga	gtacaacaga	ctcttgctcg	ctaacgacct	catgtctatc	240
aagttggacg	aatccgtgtc	cgagtctgac	accatccgga	gcatacagcat	tgtttcgcag	300
tgccctaccg	cggggaactc	ttgcctcgtn	tctggctggg	gtctgctggc	gaacggcaga	360
atgcctaccg	tgttgcactg	cgtgaacgtg	tgggtgggtg	ctgaggangt	ctgcagtaag	420
ctctatgacc	cgctgtacca	cccagcatg	ttctgcgcgc	gcggaggggc	agaccagaag	480
gactcctgca	acggtgactc	tggggggccc	ctgatctgca	acgggtactt	gcagggcctt	540
gtgtctttcg	gaaaagcccc	gtgtggccaa	cttggcgtgc	caggtgtcta	caccaacctc	600
tgcaaattca	ctgagtggat	agagaaaacc	gtccagncca	gttaactctg	gggactggga	660
acccatgaaa	ttgaccccc	aatacatcct	gcggaangaa	ttcaggaata	tctgttccca	720
gcccctctc	cctcaggccc	aggagtccag	gccccagcc	cctcctccct	caaaccaagg	780
gtacagatcc	ccagcccctc	ctccctcaga	cccaggagtc	cagaccccc	agccccctnt	840
ccntcagacc	caggagtcca	gcccctctc	cntcagacgc	aggagtccag	accccccagc	900
ccntcntccg	tcagaccacg	gggtgcaggc	ccccaacccc	tntcctntca	gagtcagagg	960
tccaagcccc	caacccctcg	ttcccagac	ccagaggtnc	aggteccagc	ccctcctccc	1020
tcagaccacg	cgggtccaatg	ccacctagan	tntcctgtga	cacagtgcgc	ccttgtggca	1080
ngttgaccca	accttaccag	ttgggtttttc	attttttgtc	cctttccccc	agatccagaa	1140
ataaagtnta	agagaagcgc	aaaaaaa				1167

<223> Xaa = Any Amino Acid

Met	Glu	Asn	Glu	Leu	Phe	Cys	Ser	Gly	Val	Leu	Val	His	Pro	Gln	Trp
1				5					10					15	
Val	Leu	Ser	Ala	Ala	His	Cys	Phe	Gln	Asn	Ser	Tyr	Thr	Ile	Gly	Leu
			20					25					30		
Gly	Leu	His	Ser	Leu	Glu	Ala	Asp	Gln	Glu	Pro	Gly	Ser	Gln	Met	Val
		35					40					45			
Glu	Ala	Ser	Leu	Ser	Val	Arg	His	Pro	Glu	Tyr	Asn	Arg	Leu	Leu	Leu
	50					55					60				
Ala	Asn	Asp	Leu	Met	Leu	Ile	Lys	Leu	Asp	Glu	Ser	Val	Ser	Glu	Ser
65					70					75					80
Asp	Thr	Ile	Arg	Ser	Ile	Ser	Ile	Ala	Ser	Gln	Cys	Pro	Thr	Ala	Gly
				85					90					95	
Asn	Ser	Cys	Leu	Val	Ser	Gly	Trp	Gly	Leu	Leu	Ala	Asn	Gly	Arg	Met
			100					105					110		
Pro	Thr	Val	Leu	His	Cys	Val	Asn	Val	Ser	Val	Val	Ser	Glu	Xaa	Val
		115					120					125			
Cys	Ser	Lys	Leu	Tyr	Asp	Pro	Leu	Tyr	His	Pro	Ser	Met	Phe	Cys	Ala



```
<210> 177
<211> 1119
<212> DNA
<213> Homo sapien
```

```
<210> 178
<211> 164
<212> PRT
<213> Homo sapien
```

<400> 178															
Met	Glu	Asn	Glu	Leu	Phe	Cys	Ser	Gly	Val	Leu	Val	His	Pro	Gln	Trp
1				5					10					15	
Val	Leu	Ser	Ala	Ala	His	Cys	Phe	Gln	Asn	Ser	Tyr	Thr	Ile	Gly	Leu
			20					25					30		







```
<210> 182
<211> 479
<212> DNA
<213> Homo sapien
```

<400>	182						
grwttk	grggatgcta	agscgccrga	rwtygtttga	tccaaccctg	gcttwttttc		60
ggaaa	atggggccta	gaagttacag	mscatytagy	tggtgcgmtg	gcacccttg		120
acag	astcccgagt	agctgggact	acaggcacac	agtcactgaa	gcaggccctg		180
aattc	acgttgccac	ctccaactta	aacattcttc	atatgtgatg	tccttagtca		240
gttaa	actttccac	ccagaaaagg	caacttagat	aaaatcttag	agtactttca		300
ctcta	agtctcttc	cagcctcact	kkgagtccm	cytgggggtt	gataggaant		360
ctggc	tttctcaata	aartctctat	ycatctcatg	tttaatttgg	tacgcata		420
cgara	aaattaaaat	gttctggtty	mactttaaaa	araaaaaaaa	aaaaaaaaa		479

<400>	183						
ggagc	agaagctaaa	gccaaagccc	aagaagagtg	gcagtgccag	cactggtgcc		60
cagta	ccaataacag	tgccagtgcc	agtgccagca	ccagtggtg	cttcagtgct		120
cagcc	tgaccgccac	tctcacattt	gggctcttcg	ctggccttg	tggagctgg		180
cacca	gtggcagctc	tggtgcctgt	ggtttctcct	acaagtgaga	ttttagatat		240
atcct	gccagtcttt	ctcttcaagc	caggggtgcat	cctcagaaac	ctactcaaca		300
ctcta	ggcagccact	atcaatcaat	tgaagttgac	actctgcatt	aratctattt		360
ctcaa	aaaaaaaaaa	aaaa					384

```
<210> 184
<211> 496
<212> DNA
<213> Homo sapien
```



<220>  
 <221> misc\_feature  
 <222> (1)...(496)  
 <223> n = A,T,C or G

<400> 184  
 accgaattgg gaccgctggc ttataagcga tcatgtyynt ccrgtatkac ctcaacgagc 60  
 agggagatcg agtctatacg ctgaagaaat ttgacccgat gggacaacag acctgctcag 120  
 cccatcctgc tcggttctcc ccagatgaca aatactctsg acaccgaatc accatcaaga 180  
 aacgcttcaa ggtgctcatg acccagcaac cgcgcctgt cctctgaggg tcccttaaac 240  
 tgatgtcttt tctgccacct gttacccctc ggagactccg taaccaaact cttcggactg 300  
 tgagccctga tgcccttttg ccagccatac tctttggcat ccagtctctc gtggcgattg 360  
 attatgcttg tgtgaggcaa tcatggtggc atcaccata aagggaacac atttgacttt 420  
 tttttctcat attttaaatt actacmagaw tattwmagaw waaatgawtt gaaaaactst 480  
 taaaaaaaaa aaaaaa 496

<210> 185  
 <211> 384  
 <212> DNA  
 <213> Homo sapien

<400> 185  
 gctggtagcc tatggcgkgg cccacggagg ggctcctgag gccacggrac agtgacttcc 60  
 caagtatcyt gcgcs gcgctc ttctaccgtc cctacctgca gatcttcggg cagattcccc 120  
 aggaggacat ggacgtggcc ctcatggagc acagcaactg ytcgtcggag cccggcttct 180  
 gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc aactggctgg 240  
 tgggtgctgt cctcgtcatc ttctgctcgt tggccaacat cctgctggtc aacttgctca 300  
 ttgccatgtt cagttacaca ttcggcгааг tacagggcaa cagcgtatctc tactgggaag 360  
 gcgcagcgtt accgcctcat ccgg 384

<210> 186  
 <211> 577  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(577)  
 <223> n = A,T,C or G

<400> 186  
 gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggcctctcgc ttcataccgc 60  
 tnccatcgtc atactgtagg tttgccacca cytcctggca tcttggggcg gcntaatatt 120  
 ccaggaaact ctcaatcaag tcaccgtcga tgaaacctgt gggctgggtc tgtcttcgcg 180  
 tcggtgtgaa aggatctccc agaaggagtg ctcgatcttc cccacacttt tgatgacttt 240  
 attgagtcga ttctgcatgt ccagcaggag gttgtaccag ctctctgaca gtgaggtcac 300  
 cagccctatc atgccgttga mcgtgccgaa garcaccgag ccttggtgtgg gggkkgaaat 360  
 ctcaccacga ttctgcatta ccagagagcc gtggcaaaag acattgacaa actcgcccag 420  
 gtggaaaaag amcamctcct ggargtgctn gccgctctc gtcmgttggt ggcagcgctw 480  
 tccttttgac acacaaacaa gttaaaggca ttttcagccc ccagaaantt gtcacatcc 540



577

```
<220>  
<221> misc_feature  
<222> (1)...(534)  
<223> n = A,T,C or G
```

```
<210> 188
<211> 761
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(761)
<223> n = A,T,C or G
```

<210>	189
<211>	482
<212>	DNA



<223> n = A, T, C or G

tttttttttt	tttgccgatn	ctactatttt	attgcaggan	gtgggggtgt	atgcaccgca	60
caccgggggt	atnagaagca	agaaggaagg	agggagggca	cagcccttg	ctgagcaaca	120
aagccgcctg	ctgccttttc	tgtctgtctc	ctgggtgcagg	cacatgggga	gaccttcccc	180
aaggcagggg	ccaccagtc	aggggtggga	atacaggggg	tgggagtgt	gcataagaag	240
tgataggcac	aggccaccgc	gtacagaccc	ctcggtcct	gacaggtnga	tttcgaccag	300
gtcattgtgc	cctgccccagg	cacagcgtn	atctggaaaa	gacagaatgc	tttccttttc	360
aaatttggt	ngtcatngaa	ngggcanttt	tccaanttng	gctnggtctt	ggtacncttg	420
gttcggccca	gctccncgtc	caaaaantat	tcaccennct	ccnaattgct	tgcnngnccc	480
cc						482

<213> Homo sapien

<223> n = A, T, C or G

tttttttttt	ttttaaaaca	gtttttcaca	acaaaattta	ttagaagaat	agtggttttg	60
aaaactctcg	catccagtga	gaactaccat	acaccacatt	acagctngga	atgtntctca	120
aatgtctggg	caaatgatac	aatggaacca	ttcaatctta	cacatgcacg	aaagaacaag	180
cgcttttgac	atacaatgca	caaaaaaaaa	aggggggggg	gaccacatgg	attaaaattt	240
taagtactca	tcacatacat	taagacacag	ttctagtcca	gtcnaaaatc	agaactgcnt	300
tgaaaaattt	catgtatgca	atccaaccaa	agaacttnat	tggtgatcat	gantnctcta	360
ctacatcnac	cttgatcatt	gccaggaacn	aaaagttnaa	ancacncngt	acaaaaanaa	420
tctgtaattn	anttcaacct	cgtacngaa	aaatnttntt	tatacactcc	c	471

<213> Homo sapien

<223> n = A, T, C or G

gagggattga aggtctgttc tastgtcggm ctgttcagcc accaactcta acaagttgct 60  
gtcttccact cactgtctgt aagcttttta acccagacwg tatcttccata aatagaacaa 120







<210> 194  
 <211> 392  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1) ... (392)  
 <223> n = A,T,C or G

<400> 194  
 gaacggctgg accttgccctc gcattgtgct tgetggcagg gaataccttg gcaagcagyt 60  
 ccagtccgag cagccccaga ccgctgccgc ccgaagctaa gcctgcctct ggccctcccc 120  
 tccgcctcaa tgcagaacca gtagtgggag cactgtgttt agagttaaga gtgaacactg 180  
 tttgatttta cttgggaatt tcctctgtta tatagctttt cccaatgcta atttccaaac 240  
 aacaacaaca aaataacatg tttgcctgtt aagttgtata aaagtaggtg attctgtatt 300  
 taaagaaaat attactgtta catatactgc ttgcaatttc tgtatttatt gktnctstgg 360  
 aaataaatat agttattaaa ggttgtcant cc 392

<210> 195  
 <211> 502  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1) ... (502)  
 <223> n = A,T,C or G

<400> 195  
 ccsttkgagg ggtkaggkyc cagttyccga gtggaagaaa caggccagga gaagtgcgtg 60  
 ccgagctgag gcagatgttc ccacagtgc cccagagcc stgggstata gtytctgacc 120  
 cctcncaagg aaagaccacs ttctggggac atgggctgga gggcaggacc tagaggcacc 180  
 aagggaaggc cccattccgg ggstgttccc cgaggaggaa ggggaaggggc tctgtgtgcc 240  
 ccccasgagg aagaggccct ggtcctggg atcagacacc ccttcacgtg tatccccaca 300  
 caaatgcaag ctcaccaagg tcccctctca gtccccttcc stacacctg amcgccact 360  
 gscscacacc caccagagc acgccaccgc ccatggggar tgtgtcaag gartcgngg 420  
 gcarcgtgga catctngtcc cagaaggggg cagaatctcc aatagangga ctgarcmtt 480  
 gctnanaaaa aaaaanaaaa aa 502

<210> 196  
 <211> 665  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1) ... (665)  
 <223> n = A,T,C or G







```

gagttgtggc tttatgttta ctgaaagtca atgcagttcc tgtacaaaga gatggccgta 360
agcattctag tacctctact ccatgggttaa gaatcgtaca cttatgttta catatgtnca 420
gggtaagaat tgtgttaagt naanttatgg agaggtccan gagaaaaatt tgatncaa 478

```

```

<210> 199
<211> 482
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

```

```

<400> 199
agtgacttgt cctccaacaa aaccccttga tcaagtttgt ggcactgaca atcagaccta 60
tgctagttcc tgctatctat tcgtacttaa atgcagactg gaggggacca aaaaggggca 120
tcaactccag ctggattatt ttggagcctg caaatctatt cctacttgta cggactttga 180
agtgattcag tttcctctac ggatgagaga ctggctcaag aatatacctca tgcagcttta 240
tgaagccnac tctgaacacg ctggttatct nagatgagaa ncagagaaat aaagtcnaga 300
aaatttacct ggangaaaag aggccttngg ctggggacca tcccattgaa ccttctctta 360
anggacttta agaanaaaact accacatgtn tgtngtatcc tgggtgccngg ccgtttantg 420
aacntngacn ncacccttnt ggaatanant cttgacngcn tcctgaactt gctcctctgc 480
ga 482

```

```

<210> 200
<211> 270
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(270)
<223> n = A,T,C or G

```

```

<400> 200
cggccgcaag tgcaactcca gctggggcgc tgccgacgaa gattctgcca gcagttggtc 60
cgactgcgac gacggcggcg gcgacagtcg caggtgcagc gcgggcgcct ggggtcttgc 120
aaggctgagc tgacgccgca gaggtcgtgt cacgtcccac gaccttgacg ccgtcgggga 180
cagccggaac agagcccggg gaangcggga ggcctcgggg agccctcggg gaagggcggc 240
ccgagagata cgcaggtgca ggtggccgcc 270

```

```

<210> 201
<211> 419
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(419)
<223> n = A,T,C or G

```



```

<400> 201
tttttttttt ttttgaatc tactgcgagc acagcaggtc agcaacaagt ttattttgca      60
gctagcaagg taacagggta gggcatggtt acatgttcag gtcaacttcc tttgtcgtgg      120
ttgattgggt tgtctttatg ggggcggggt ggggtagggg aaancgaagc anaantaaca      180
tggagtgggt gcacctccc tgtagaacct gggtacnaaa gcttggggca gttcacctgg      240
tctgtgaccg tcattttctt gacatcaatg ttattagaag tcaggatata ttttagagag      300
tccactgtnt ctggaggagg attaggggtt cttgccana tccaancaa atccacntga      360
aaaagttgga tgatncangt acngaatacc ganggcatan ttctcatant cggtggccca      419

```

```

<210> 202
<211> 509
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (509)
<223> n = A,T,C or G

```

```

<400> 202
tttntttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt      60
tggcacttaa tccattttta tttcaaaatg tctacaaant ttnaatncnc cattatacng      120
gtnatnttnc aaaatctaaa nnttattcaa atntnagcca aantccttac ncaaattnaa      180
tacnncnaaa aatcaaaaat atacntntct ttcagcaaac ttngttacat aaattaaaaa      240
aatatatacg gctgggtgtt tcaaagtaca attatcttaa cactgcaaac atnttttnaa      300
ggaactaaaa taataaaaaa cactnccgca aagggttaaag ggaacaacaa attcntttta      360
caacancnnc nattataaaa atcatatctc aaatccttag ggaatatata cttcacacng      420
ggatcttaac ttttactnca ctttgtttat ttttttanaa ccattgtntt gggcccaaca      480
caatggnaat nccnccnccn tggactagt                                     509

```

```

<210> 203
<211> 583
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (583)
<223> n = A,T,C or G

```

```

<400> 203
tttttttttt ttttttttga cccccctctt ataaaaaaca agttaccatt ttattttact      60
tacacatatt tattttataa ttggtattag atattcaaaa ggcagctttt aaaatcaaac      120
taaattgaaa ctgccttaga tacataattc ttaggaatta gcttaaaaac tgcctaaagt      180
gaaaatcttc tctagctctt ttgactgtaa atttttgact cttgtaaaac atccaaattc      240
atttttcttg tctttaaaat tatctaactt ttccattttt tccctattcc aagtcaattt      300
gcttctctag cctcatttcc tagctottat ctactattag taagtggctt ttttcctaaa      360
agggaaaaca ggaagagana atggcacaca aaacaaacat tttatattca tatttctacc      420
tacgttaata aaatagcatt ttgtgaagcc agctcaaaaag aaggcttaga tccttttatg      480
tccattttag tcactaaacg atatcnaaag tgccagaatg caaaagggtt gtgaacattt      540

```



583

```
<220>  
<221> misc_feature  
<222> (1)...(589)  
<223> n = A,T,C or G
```

```
<210> 205
<211> 545
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(545)
<223> n = A,T,C or G
```

```
<210> 206
<211> 487
<212> DNA
<213> Homo sapien
```



<220>  
 <221> misc\_feature  
 <222> (1)...(487)  
 <223> n = A,T,C or G

<400> 206  
 tttttttttt ttttttagtc aagtttctna tttttattat aattaaagtc ttgggtcattt 60  
 catttatttag ctctgcaact tacatattta aattaaagaa acgttnttag acaactgtna 120  
 caatttataa atgtaagggtg ccattattga gtanatatat tcctccaaga gtggatgtgt 180  
 cccttctccc accaactaat gaancagcaa cattagttta attttattag tagatnatac 240  
 actgctgcaa acgctaattc tcttctccat ccccatgtng atattgtgta tatgtgtgag 300  
 ttggttagaa tgcatacanca atctnacaat caacagcaag atgaagctag gcntgggctt 360  
 tcggtgaaaa tagactgtgt ctgtctgaat caaatgatct gacctatcct cgggtggcaag 420  
 aactcttcga accgcttcct caaaggcngc tgccacattt gtggcntctn ttgcacttgt 480  
 ttcaaaa 487

<210> 207  
 <211> 332  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(332)  
 <223> n = A,T,C or G

<400> 207  
 tgaattggct aaaagactgc atttttanaa ctagcaactc ttattttcttt cttttaaaaa 60  
 tacatagcat taaatcccaa atcctattta aagacctgac agcttgagaa ggctcactact 120  
 gcatttatag gaccttctgg tggttctgct gttacntttg aantctgaca atccttgana 180  
 atctttgcat gcagaggagg taaaaggat tggattttca cagaggaana acacagcgca 240  
 gaaatgaagg ggccaggctt actgagcttg tccactggag ggctcatggg tgggacatgg 300  
 aaaagaaggc agcctaggcc ctggggagcc ca 332

<210> 208  
 <211> 524  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(524)  
 <223> n = A,T,C or G

<400> 208  
 agggcgtggt gcgaggggcg ttactgtttt gtctcagtaa caataaatac aaaaagactg 60  
 gttgtgttcc ggccccatcc aaccacgaag ttgatttctc ttgtgtgcag agtgactgat 120  
 tttaaaggac atggagcttg tcacaatgtc acaatgtcac agtgtgaagg gcacactcac 180  
 tcccgcgtga ttcacattta gcaaccaaca atagctcatg agtccatact tgtaaatact 240  
 tttggcagaa tacttnttga aacttgcaga tgataactaa gatccaagat atttcccaa 300  
 gtaaatagaa gtgggtcata atattaatta cctgttcaca tcagcttcca tttacaagtc 360



atgagcccag	acactgacat	caaactaagc	ccacttagac	tcctcaccac	cagtctgtcc	420
tgtcatcaga	caggaggctg	tcaccttgac	caaattctca	ccagtcaatc	atctatccaa	480
aaaccattac	ctgatccact	tcgggtaatg	caccaccttg	gtga		524

<210> 209  
 <211> 159  
 <212> DNA  
 <213> Homo sapien

<400> 209						
gggtgaggaa	atccagagtt	gccatggaga	aaattccagt	gtcagcattc	ttgtctcttg	60
tggccctctc	ctacactctg	gccagagata	ccacagtcaa	acctggagcc	aaaaaggaca	120
caaaggactc	tcgaccctaa	ctgccccaga	ccctctcca			159

<210> 210  
 <211> 256  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(256)  
 <223> n = A,T,C or G

<400> 210						
actccctggc	agacaaaggc	agaggagaga	gctctgttag	ttctgtgttg	ttgaactgcc	60
actgaatttc	tttccacttg	gactattaca	tgccanttga	gggactaatg	gaaaaacgta	120
tggggagatt	ttanccaatt	tangtntgta	aatggggaga	ctggggcagg	cgggagagat	180
ttgcagggtg	naaatgggan	ggctggtttg	ttanatgaac	agggacatag	gaggtaggca	240
ccaggatgct	aatca					256

<210> 211  
 <211> 264  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(264)  
 <223> n = A,T,C or G

<400> 211						
acattgtttt	tttgagataa	agcattgaga	gagctctcct	taacgtgaca	caatggaagg	60
actggaacac	ataccacat	ctttgttctg	agggataatt	ttctgataaa	gtcttgctgt	120
atattcaagc	acatatgtta	tatattattc	agttccatgt	ttatagccta	gttaaggaga	180
ggggagatac	attcngaaag	aggactgaaa	gaaatactca	agtnngaaaa	cagaaaaaga	240
aaaaaaggag	caaatgagaa	gcct				264

<210> 212  
 <211> 328  
 <212> DNA



<400> 214						
accagaatc	caatgctgaa	tatttggtt	cattattccc	agattctttg	attgtcaaag	60
gatttaatgt	tgtctcagct	tgggcacttc	agttaggacc	taaggatgcc	agccggcagg	120
tttatatatg	cagcaacaat	attcaagcgc	gacaacaggt	tattgaactt	gcccgccagt	180
tgaatttcat	tcccatggac	ttgggatcct	tatcatcagc	canagagatt	gaaaatttac	240
ccctacgact	ctttactctc	tggagagggc	cagtgggtgt	agctataagc	ttggccacat	300
ttttttttcc	tttattcctt	tgtcagagat	gcgattcatc	catatgctan	aaaccaacag	360
agtgactttt	acaaaattcc	tataganatt	gtgaataaaa	ccttacctat	agttgccatt	420
actttgctct	ccctaataata	cctc				444



<400> 215

<400> 216

<400> 217

acctacgtgg	gtaagtttan	aaatgttata	atttcaggaa	naggaacgca	tataattgta	60
tcttgctat	aattttctat	tttaataagg	aaatagcaaa	ttgggtggg	gggaatgtag	120
ggcattctac	agtttgagca	aaatgcaatt	aaatgtggaa	ggacagcact	gaaaaatttt	180



atgaataatc tgtatgatta tatgtctcta gagtagattt ataattagcc acttacccta 240  
atataccttca tgcttgtaaa gt 262

<210> 218  
<211> 205  
<212> DNA  
<213> Homo sapien  
  
<220>  
<221> misc\_feature  
<222> (1)...(205)  
<223> n = A,T,C or G

<400> 218  
accaaggtgg tgcattaccg gaantggatc aangacacca tcgtggccaa cccctgagca 60  
cccctatcaa ctcccttttg tagtaaaactt ggaaccttgg aaatgaccag gccaaagactc 120  
aggcctcccc agttctactg acctttgtcc ttangtntna ngcccagggt tgctaggaaa 180  
anaaatcagc agacacaggt gtaaa 205

<210> 219  
<211> 114  
<212> DNA  
<213> Homo sapien

<400> 219  
tactgttttg tctcagtaac aataaatata aaaagactgg ttgtgttccg gccccatcca 60  
accacgaagt tgatttctct tgtgtgcaga gtgactgatt ttaaaggaca tgga 114

<210> 220  
<211> 93  
<212> DNA  
<213> Homo sapien

<400> 220  
actagccagc acaaaaaggca gggtagcctg aattgctttc tgctctttac atttctttta 60  
aaataagcat ttagtgctca gtcccactg agt 93

<210> 221  
<211> 167  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(167)  
<223> n = A,T,C or G

<400> 221  
actangtgca ggtgcgaca aatatttgct gatattccct tcactcttga ttccatgagg 60  
tcttttgccc agcctgtggc tctactgtag taagtttctg ctgatgagga gccagnatgc 120  
ccccactac ctccctgac gctccccana aatcacccaa cctctgt 167



<210> 222  
 <211> 351  
 <212> DNA  
 <213> Homo sapien

<400> 222  
 agggcggtggt gcggagggcg gtactgacct cattagtagg aggatgcatt ctggcacccc 60  
 gttcttcacc tgtcccccaa tccttaaaag gccatactgc ataaagtcaa caacagataa 120  
 atgtttgctg aattaaagga tggatgaaaa aaattaataa tgaatttttg cataatccaa 180  
 ttttctcttt tatatttcta gaagaagttt ctttgagcct attagatccc gggaatcttt 240  
 taggtgagca tgattagaga gctttagggt tgcttttaca tatatctggc atatttgagt 300  
 ctcgtatcaa aacaatagat tggtaaagggt ggtattattg tattgataag t 351

<210> 223  
 <211> 383  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(383)  
 <223> n = A,T,C or G

<400> 223  
 aaaacaaaca acaaaaaaaa acaattcttc attcagaaaa attatcttag ggactgatat 60  
 tggtaattat ggtcaattta atwrtrttkt ggggcatttc cttacattgt cttgacaaga 120  
 ttaaaatgtc tgtgccaaaa ttttgtattt tatttggaaga cttcttatca aaagtaatgc 180  
 tgccaaagga agtctaagga attagtagtg ttcccmctac ttgtttggag tgtgctattc 240  
 taaaagattt tgatttcctg gaatgacaat tatattttta ctttggtggg ggaaanagtt 300  
 ataggaccac agtcttcact tctgatactt gtaaattaat cttttattgc acttgttttg 360  
 accattaagc tatatgttta aaa 383

<210> 224  
 <211> 320  
 <212> DNA  
 <213> Homo sapien

<400> 224  
 cccctgaagg cttcttggtta gaaaatagta cagttacaac caataggaac aacaaaaaga 60  
 aaaagtttgt gacattgtag tagggagtgt gtacccttca ctcccatca aaaaaaaat 120  
 ggatacatgg ttaaaggata raagggcaat attttatcat atgttctaaa agagaaggaa 180  
 gagaaaatac tactttctcr aaatggaagc ccttaaagggt gctttgatac tgaaggacac 240  
 aaatgtggcc gtccatcctc ctttaragtt gcatgacttg gacacggtaa ctgttgacgt 300  
 tttaractcm gcattgtgac 320

<210> 225  
 <211> 1214  
 <212> DNA  
 <213> Homo sapien



&lt;400&gt; 225

gaggactgca	gcccgcactc	gcagccctgg	caggcgccac	tggatcatgga	aaacgaattg	60
ttctgctcgg	gcgtcctggg	gcatccgcag	tgggtgctgt	cagccgcaca	ctgtttccag	120
aactcctaca	ccatcgggct	gggcctgcac	agtcttgagg	ccgaccaaga	gccagggagc	180
cagatgggtg	aggccagcct	ctccgtacgg	cacccagagt	acaacagacc	cttgctcgtc	240
aacgacctca	tgctcatcaa	gttggacgaa	tccgtgtccg	agtctgacac	catccggagc	300
atcagcattg	cttcgcagtg	ccctaccgcg	gggaactctt	gcctcgtttc	tggctggggg	360
ctgctggcga	acggcagaat	gcctaccgtg	ctgcagtgcg	tgaacgtgtc	gggtggtgtc	420
gaggaggtct	gcagtaagct	ctatgacccg	ctgtaccacc	ccagcatgtt	ctgcgcgggc	480
ggagggcaag	accagaagga	ctcctgcaac	ggtgactctg	ggggggccct	gatctgcaac	540
gggtacttgc	agggccttgt	gtcttttcgga	aaagccctgt	gtggccaagt	tggcgtgcc	600
gggtgtctaca	ccacctctg	caaattcact	gagtggatag	agaaaaccgt	ccaggccagt	660
taactctggg	gactgggaac	ccatgaaatt	gacccccaaa	tacatcctgc	ggaaggaatt	720
caggaatata	tgttcccagc	ccctcctccc	tcaggcccag	gagtccaggc	ccccagcccc	780
tcctccctca	aaccaagggt	acagatcccc	agccctcctc	ccctcagacc	caggagtcca	840
gacccccccag	ccctcctccc	ctcagaccca	ggagtccagc	ccctcctccc	tcagacccag	900
gagtccagac	ccccagcccc	ctcctccctc	agaccagggg	gtccaggccc	ccaacccctc	960
ctccctcaga	ctcagagggt	caagccccca	acccctcctt	cccagacccc	agaggtccag	1020
gtcccagccc	ctcctccctc	agacccagcg	gtccaatgcc	acctagaact	tcctgttaca	1080
cagtgcctcc	ttgtggcag	ttgacccaac	cttaccagtt	ggtttttcat	ttttgttccc	1140
tttcccttag	atccagaaat	aaagtctaag	agaagcgcaa	aaaaaaaaaa	aaaaaaaaaa	1200
aaaaaaaaaa	aaaa					1214

&lt;210&gt; 226

&lt;211&gt; 119

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 226

accagtatg	tgcagggaga	cggaacccca	tgtgacagcc	cactccacca	gggttcccaa	60
agaacctggc	ccagtcataa	tcattcatcc	tgacagtggc	aataatcacg	ataaccagt	119

&lt;210&gt; 227

&lt;211&gt; 818

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 227

acaattcata	gggacgacca	atgaggacag	ggaatgaacc	cggtctctcc	ccagccctga	60
tttttgctac	atatgggggc	ccttttcatt	ctttgcaaaa	acactggggt	ttctgagaac	120
acggacgggt	cttagcacia	tttgtgaaat	ctgtgtaraa	ccgggctttg	caggggagat	180
aattttctct	ctctggagga	aagggtggtg	ttgacaggca	gggagacagt	gacaaggcta	240
gagaaagcca	cgctcggcct	tctctgaacc	aggatggaac	ggcagacccc	tgaaaacgaa	300
gcttgtcccc	ttccaatcag	ccaattctga	gaacccccat	ctaacttctc	actggaaaag	360
agggcctcct	caggagcagt	ccaagagttt	tcaaagataa	cgtgacaact	accatctaga	420
ggaaagggtg	caccctcagc	agagaagccg	agagcttaac	tctggctcgt	tccagagaca	480
acctgctggc	tgtcttgagg	tgcgcccagc	ctttgagagg	ccactacccc	atgaacttct	540
gccatccact	ggacatgaag	ctgaggacac	tgggcttcaa	cactgagttg	tcatgagagg	600
gacaggctct	gccctcaagc	cggtgagggg	cagcaaccac	tctcctcccc	tttctcacgc	660
aaagccattc	ccacaaatcc	agaccatacc	atgaagcaac	gagacccaaa	cagtttggct	720
caagaggata	tgaggactgt	ctcagcctgg	ctttgggctg	acaccatgca	cacacacaag	780



818

<400> 228

```
<210> 229
<211> 300
<212> DNA
<213> Homo sapien
```

<400> 229

```
<210> 230
<211> 301
<212> DNA
<213> Homo sapien
```

<400> 230

```
<210> 231
<211> 301
<212> DNA
<213> Homo sapien
```



```
<210> 235
<211> 283
<212> DNA
```



<400> 235

<210> 236

<211> 301

<212> DNA

<213> Homo sapien

<400> 236

<210> 237

<211> 301

<212> DNA

<213> Homo sapien

<400> 237

<210> 238

<211> 301

<212> DNA

<213> Homo sapien

<400> 238

<210> 239

<211> 239

<212> DNA



<213> Homo sapien

<400> 239

ataagcagct	agggaattct	ttatttagta	atgtcctaac	ataaaagttc	acataactgc	60
ttctgtcaaa	ccatgatact	gagctttgtg	acaaccaga	aataactaag	agaaggcaaa	120
cataatacct	tagagatcaa	gaaacattta	cacagttcaa	ctgttttaaaa	atagctcaac	180
attcagccag	tgagtagagt	gtgaatgcc	gcatacacag	tatacaggtc	cttcaggga	239

<210> 240

<211> 300

<212> DNA

<213> Homo sapien

<400> 240

ggtcctaattg	aagcagcagc	ttccacattt	taacgcaggt	ttacgggtgat	actgtccttt	60
gggatctgcc	ctccagtgg	acctttttaag	gaagaagtgg	gccaagcta	agttccacat	120
gctgggtgag	ccagatgact	tctgttccct	ggtcactttc	ttcaatgggg	cgaatggggg	180
ctgccagggt	tttaaaatca	tgtttcatct	tgaagcacac	ggtcacttca	ccctctcac	240
gctgtgggtg	tactttgatg	aaaataccca	ctttgttggc	ctttctgaag	ctataatgtc	300

<210> 241

<211> 301

<212> DNA

<213> Homo sapien

<400> 241

gaggtctggt	gctgaggtct	ctgggctagg	aagaggagtt	ctgtggagct	ggaagccaga	60
cctcttttga	ggaaactcca	gcagctatgt	tgggtgtctct	gagggaatgc	aacaaggctg	120
ctcctccatg	tattggaaaa	ctgcaaaactg	gactcaactg	gaagggaagtg	ctgctgccag	180
tgtgaagaac	cagcctgagg	tgacagaaac	ggaagcaaac	aggaacagcc	agtcttttct	240
tctcctcct	gtcatacgg	ctctctcaag	catcctttgt	tgtcaggggc	ctaaaaggga	300
g						301

<210> 242

<211> 301

<212> DNA

<213> Homo sapien

<400> 242

ccgaggtcct	gggatgcaac	caatcaactct	gtttcacgtg	acttttatca	ccatacaatt	60
tgtggcattt	cctcattttc	tacattgtag	aatcaagagt	gtaaataaat	gtatatcgat	120
gtcttcaaga	atatatcatt	cctttttcac	tagaaccat	tcaaaatata	agtcaagaat	180
cttaatatca	acaaatatat	caagcaaact	ggaaggcaga	ataactacca	taatttagta	240
taagtacca	aagttttata	aatcaaaagc	cctaatagata	accattttta	gaattcaatc	300
a						301

<210> 243

<211> 301

<212> DNA

<213> Homo sapien



<213> Homo sapien



&lt;400&gt; 247

```

aggtcctttg gcagggctca tggatcagag ctcaaactgg agggaaaggc atttcgggta      60
gcctaagagg gcgactggcg gcagcacaac caaggaaggc aaggttgttt cccccacgct      120
gtgtcctgtg ttcaggtgcg acacacaatc ctcattgggaa caggatcacc catgcgctgc      180
ccttgatgat caaggttggg gcttaagtgg attaaggag gcaagttctg gggttccttg      240
cttttcaaac catgaagtca ggctctgtat ccctcctttt cctaactgat attctaacta      300
a                                                                                   301

```

&lt;210&gt; 248

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 248

```

aggtccttgg agatgccatt tcagccgaag gactcttctw ttcggaagta caccctcact      60
attaggaaga ttcttagggg taatTTTTTct gaggaaggag aactagccaa cttaagaatt      120
acaggaagaa agtgggttgg aagacagcca aagaaataaa agcagattaa attgtatcag      180
gtacattcca gcctgttggc aactccataa aaacatttca gatttttaatc ccgaatttag      240
ctaattgagac tggatttttg ttttttatgt tgtgtgtcgc agagctaaaa actcagttcc      300
c                                                                                   301

```

&lt;210&gt; 249

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 249

```

gtccagagga agcacctggg gctgaactag gcttgccctg ctgtgaactt gcacttggag      60
ccctgacgct gctgttctcc ccgaaaaacc cgaccgacct ccgcgatctc cgtcccgcgc      120
ccagggagac acagcagtga ctacagagctg gtcgcacact gtgcctccct cctcaccgcc      180
catcgtaatg aattattttg aaaattaatt ccaccatcct ttcagattct ggatggaaaag      240
actgaatctt tgactcagaa ttgtttgctg aaaagaatga tgtgactttc ttagtcattt      300
a                                                                                   301

```

&lt;210&gt; 250

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 250

```

ggctctgtgac aaggacttgc aggetgtggg aggcaagtga cccttaacac tacacttctc      60
cttatcttta ttggcttgat aaacataatt atttctaaca ctagcttatt tccagttgcc      120
cataagcaca tcagtacttt tctctggctg gaatagtaaa cttaaagtatg gtacatctac      180
ctaaaagact actatgtgga ataatacata ctaatgaagt attacatgat ttaaagacta      240
caataaaacc aaacatgctt ataacattaa gaaaaacaat aaagatacat gattgaaacc      300
a                                                                                   301

```

&lt;210&gt; 251

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien



```
<210> 252
<211> 301
<212> DNA
<213> Homo sapien
```

```
<210> 253
<211> 301
<212> DNA
<213> Homo sapien
```

```
<210> 254
<211> 301
<212> DNA
<213> Homo sapien
```

<210>	255
<211>	302
<212>	DNA



<213> Homo sapien

<400> 255

```
agcttttttt tttttttttt tttttttttt ttcattaaaa aatagtgtctc tttattataa      60
attactgaaa tgtttctttt ctgaatataa atataaatat gtgcaaagtt tgacttggat      120
tgggattttg ttgagttctt caagcatctc ctaataccct caagggcctg agtagggggg      180
aggaaaaagg actggaggtg gaatctttat aaaaaacaag agtgattgag gcagattgta      240
aacattatta aaaaacaaga aacaaacaaa aaaatagaga aaaaaaccac cccaacacac      300
aa                                                                302
```

<210> 256

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 256

```
gttccagaaa acattgaagg tggcttccca aagtctaact agggataccc cctctagcct      60
aggaccctcc tccccacacc tcaatccacc aaaccatcca taatgcaccc agataggccc      120
acccccaaaa gcctggacac cttagacaca cagttatgac caggacagac tcatctctat      180
aggcaaatag ctgctggcaa actggcatta cctggtttgt ggggatgggg gggcaagtgt      240
gtggcctctc ggccctggta gcaagaacat tcagggtagg cctaagttan tcgtgttagt      300
t                                                                301
```

<210> 257

<211> 301

<212> DNA

<213> Homo sapien

<400> 257

```
gttgtggagg aactctggct tgctcattaa gtctactga ttttactat cccctgaatt      60
tccccactta tttttgtctt tcaatatcgc aggccttaga agaggtctac ctgcctccag      120
tcttacctag tccagtctac cccctggagt tagaatggcc atcctgaagt gaaaagtaat      180
gtcacattac tcccttcagt gatttcttgt agaagtgcc atccctgaat gccaccaaga      240
tcttaatctt cacatcttta atcttatctc tttgactcct ctttacaccg gagaaggctc      300
c                                                                301
```

<210> 258

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(301)

<223> n = A,T,C or G



&lt;400&gt; 258

```

cagcagtagt agatgccgta tgccagcacg cccagcactc ccaggatcag caccagcacc      60
aggggcccag ccaccaggcg cagaagcaag ataaacagta ggctcaagac cagagccacc      120
cccagggcaa caagaatcca ataccaggac tggggcaaat cttcaaagat cttaacactg      180
atgtctcggg cattgaggct gtcaataana cgctgatccc ctgctgtatg gtggtgtcat      240
tggtgatccc tgggagcgcc ggtggagtaa cgttgggtcca tggaaagcag cgcccacaac      300
t                                                                                   301

```

&lt;210&gt; 259

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1) ... (301)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 259

```

tcatatatgc aaacaaatgc agactangcc tcaggcagag actaaaggac atctcttggg      60
gtgtcctgaa gtgatttggg cccctgaggg cagacaccta agtaggaatc ccagtgggaa      120
gcaaagccat aaggaagccc aggattcctt gtgatcagga agtgggccag gaaggctctgt      180
tccagctcac atctcatctg catgcagcac ggaccggatg cgcccactgg gtcttggctt      240
ccctcccatc ttctcaagca gtgtccttgt tgagccattt gcatccttgg ctccagggtg      300
c                                                                                   301

```

&lt;210&gt; 260

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 260

```

ttttttttct ccctaaggaa aaagaaggaa caagtctcat aaaaccaa at aagcaatggt      60
aagggtgtctt aacttgaaaa agattaggag tctctgggtt acaagttata attgaatgaa      120
agaactgtaa cagccacagt tggccatttc atgccaatgg cagcaaacia caggattaac      180
tagggcaaaa taaataagtg tgtggaagcc ctgataagtg cttaataaac agactgattc      240
actgagacat cagtacctgc ccgggcggcc gctcgagccg aattctgcag atatccatca      300
c                                                                                   301

```

&lt;210&gt; 261

&lt;211&gt; 301

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 261

```

aaatattcga gcaaatcctg taactaatgt gtctccataa aaggctttga actcagtgaa      60
tctgcttcca tccacgattc tagcaatgac ctctcggaca tcaaagctcc tottaagggt      120
agcaccaact attccatata attcatcagc aggaaataaa ggctcttcag aagggttcaat      180
ggtgacatcc aatttcttct gataatttag attcctcaca accttcctag ttaagtgaag      240
ggcatgatga tcatccaaag cccagtggtc acttactcca gactttctgc aatgaagatc      300
a                                                                                   301

```



<210> 262  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<400> 262  
 gaggagagcc tgttacagca tttgtaagca cagaatactc caggagtatt tgtaattgtc 60  
 tgtgagcttc ttgccgcaag tctctcagaa atttaaaaag atgcaaatacc ctgagtcacc 120  
 cctagacttc ctaaaccaga tcctctgggg ctggaacctg gcaactctgca tttgtaatga 180  
 gggctttctg gtgcacacct aattttgtgc atctttgccc taaatcctgg attagtgcc 240  
 catcattacc cccacattat aatgggatag attcagagca gatactctcc agcaaagaat 300  
 c 301

<210> 263  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(301)  
 <223> n = A,T,C or G

<400> 263  
 tttagcttgt ggtaaatgac tcacaaaact gattttaaaa tcaagttaat gtgaattttg 60  
 aaaattacta cttaatccta attcacaata acaatggcat taaggtttga cttgagttgg 120  
 ttcttagtat tatttatggg aaataggctc ttaccacttg caaataactg gccacatcat 180  
 taatgactga ctctccagta aggctctcta aggggtaagt angaggatcc acaggatttg 240  
 agatgctaag gccccagaga tcgtttgatc caaccctctt attttcagag gggaaaatgg 300  
 g 301

<210> 264  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<400> 264  
 aaagacgtta aaccactcta ctaccacttg tggaactctc aaagggtaaa tgacaaascc 60  
 aatgaatgac tctaaaaaca atatttacat ttaatggttt gtagacaata aaaaaacaag 120  
 gtggatagat ctagaattgt aacattttta gaaaaccata scatttgaca gatgagaaag 180  
 ctcaattata gatgcaaagt tataactaaa ctactatagt agtaaagaaa tacatttcac 240  
 acccttcata taaattcact atcttggctt gaggcactcc ataaaatgta tcacgtgcat 300  
 a 301

<210> 265  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<400> 265



```

tgcccaagtt atgtgtaagt gtatccgcac ccagaggtaa aactacactg tcatctttgt      60
cttcttgtga cgcagtatct cttctctggg gagaagccgg gaagtcttct cctggctcta      120
catattcttg gaagtctcta atcaactttt gttccatttg tttcatttct tcaggaggga      180
ttttcagttt gtcaacatgt tctctaacia cacttgccca tttctgtaaa gaatccaaag      240
cagtccaagg ctttgacatg tcaacaacca gcataactag agtatccttc agagatacgg      300
c                                                                                   301

```

```

<210> 266
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 266
taccgtctgc ccttctctcc atccaggcca tctgcgaatc tacatggggtc ctccattctg      60
acaccagatc actcttttct ctaccacag gcttgctatg agcaagagac acaacctctc      120
ctcttctgtg ttccagcttc ttttctgtt cttcccaccc cttaagttct attcctgggg      180
atagagacac caatacccat aacctctctc ctaagcctcc ttataaccca ggggtgcacag      240
cacagactcc tgacaactgg taaggccaat gaactgggag ctcacagctg gctgtgcctg      300
a                                                                                   301

```

```

<210> 267
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 267
aaagagcaca ggccagctca gcctgccctg gccatctaga ctcagcctgg ctccatgggg      60
gttctcagtg ctgagtcctc ccaggaaaag ctcacctaga ccttctgagg ctgaatcttc      120
atcctcacag gcagcttctg agagcctgat attcctagcc ttgatgggtc ggagtaaagc      180
ctcattctga ttctctctct tcttttcttt caagttgggt ttcctcacat ccctctgttc      240
aattcgcttc agcttgtctg ctttagccct catttcaga agcttcttct ctttggcctc      300
t                                                                                   301

```

```

<210> 268
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 268
aatgtctcac tcaactactt cccagcctac cgtggcctaa ttctgggagt tttcttctta      60
gatcttggga gagctgggtc ttctaaggag aaggaggaag gacagatgta actttggatc      120
tcgaagagga agtctaattg aagtaattag tcaacgggtc ttgttttagac tcttgggaata      180
tgctgggtgg ctcagtgagc ccttttggag aaagcaagta ttattcttaa ggagtaacca      240
cttccattg ttctactttc taccatcctc aattgtatat tatgtattct ttggagaact      300
a                                                                                   301

```

```

<210> 269
<211> 301
<212> DNA
<213> Homo sapien

```



taacaatata	cactagctat	ctttttaact	gtccatcatt	agcaccaatg	aagattcaat	60
aaaattacct	ttattcacac	atctcaaaac	aattctgcaa	attcttagtg	aagtttaact	120
atagtcacag	accttaaata	ttcacattgt	tttctatgtc	tactgaaaat	aagttcacta	180
cttttctgga	tattctttac	aaaatcttat	taaaattcct	ggtattatca	cccccaatta	240
tacagtagca	caaccacctt	atgtagtttt	tacatgatag	ctctgtagaa	gtttcacatc	300
t						301

<213> Homo sapien

a

<213> Homo sapien

<223> n = A, T, C or G

C

<213> Homo sapien

g



<210> 273  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(301)  
 <223> n = A,T,C or G

<400> 273  
 acatgtgtgt atgtgtatct ttgggaaan aanaagacat cttgtttayt attttttttg 60  
 agagangctg ggacatggat aatcacwtaa ttgctayta tyactttaat ctgactygaa 120  
 gaaccgtcta aaaataaaat ttaccatgtc dtatattcct tatagtatgc ttatttcacc 180  
 ttytttctgt ccagagagag tatcagtgac ananatttma gggatgaamac atgmattggg 240  
 gggacttnty tttacngagm accctgcccc sgcgcctcg makcngantt ccgcsananc 300  
 t 301

<210> 274  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(301)  
 <223> n = A,T,C or G

<400> 274  
 cttatatact ctttctcaga ggcaaaagag gagatgggta atgtagacaa ttctttgagg 60  
 aacagtaa at gattattaga gagaangaat ggaccaagga gacagaaatt aacttgtaaa 120  
 tgattctctt tggaatctga atgagatcaa gaggccagct ttagcttctg gaaaagtcca 180  
 tctaggtatg gttgcattct cgtcttcttt tctgcagtag ataatgaggt aaccgaaggc 240  
 aattgtgctt cttttgataa gaagctttct tggatcatatc aggaaattcc aganaaagtc 300  
 c 301

<210> 275  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(301)  
 <223> n = A,T,C or G

<400> 275  
 tcggtgtcag cagcacgtgg cattgaacat tgcaatgtgg agcccaaacc acagaaaatg 60  
 gggatgaa at ggccaacttt ctattaactt atgttggaac ttttgccacc aacagtaagc 120  
 tggcccttct aataaaaagaa aattgaaagg tttctcacta aacggaatta agtagtgag 180



```

<210> 276
<211> 301
<212> DNA
<213> Homo sapien

<400> 276
tgtacacata ctcaataaat aaatgactgc attgtggtat tattactata ctgattatat      60
ttatcatgtg acttctaatt agaaaatgta tccaaaagca aaacagcaga tatacaaaaat      120
taaagagaca gaagatagac attaacagat aaggcaactt atacattgag aatccaaatc      180
caatacattt aaacatttgg gaaatgaggg ggacaaatgg aagccagatc aaatttgtgt      240
aaaactattc agtatgtttc ccttgcttca tgtctgagaa ggctctcctt caatggggat      300
q                                          301

```

<400> 277						
tttgttgatg	tcagtatttt	attactttgcg	ttatgagtg	tcacctggga	aattctaaag	60
atacagagga	cttggaggaa	gcagagcaac	tgaatttaat	ttaaaagaag	gaaaacattg	120
gaatcatggc	actcctgata	ctttcccaa	tcaacactct	caatgcccc	ccctcgtoct	180
caccatagtg	gggagactaa	agtggccacg	gatttgctt	angtgtgcag	tgcgttctga	240
gttcnctgtc	gattacatct	gaccagtctc	ctttttccga	agtccntccg	ttcaatcttg	300
c						301

<400> 278							
taccactaca	ctccagcctg	ggcaacagag	caagacctgt	ctcaaagcat	aaaatggaat		60
aacatatcaa	atgaaacagg	gaaaatgaag	ctgacaattt	atggaagcca	gggcttgtca		120
cagtctctac	tgttattatg	cattacctgg	gaatttatat	aagcccttaa	taataatgcc		180
aatgaacatc	tcatgtgtgc	tcacaatgtt	ctggcactat	tataagtgtc	tcacaggttt		240
tatgtgtttc	tcgtaacttt	atggantagg	tactcggccq	cgaacacqct	aaqcqaatt		300



301

<400> 279

```
<210> 280
<211> 301
<212> DNA
<213> Homo sapien
```

<400> 280

```
<210> 281
<211> 301
<212> DNA
<213> Homo sapien
```

<400> 281

```
<210> 282
<211> 301
<212> DNA
<213> Homo sapien
```



acatcaccat	gatcggtacc	cccacccatt	atacgttgta	tgtttacata	aatactcttc	60
aatgatcatt	agtgttttaa	aaaaaatact	gaaaactcct	tctgcatccc	aatctctaac	120
caggaaagca	aatgctattt	acagacctgc	aagccctccc	tcaaacnaaa	ctatttctgg	180
attaaatatg	tctgacttct	tttgagggtca	cacgactagg	caaatgctat	ttacgatctg	240
caaaagctgt	ttgaagagtc	aaagcccca	tgtgaacacg	atttctggac	cctgtaacag	300
t						301



<210> 286  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<400> 286  
 taccactgca ttccagcctg ggtgacagag tgagactccg tctccaaaaa aaactttgct 60  
 tgtatattat ttttgcctta cagtggatca ttctagtagg aaaggacagt aagatTTTTT 120  
 atcaaaatgt gtcatgccag taagagatgt tatattcttt tctcatttct tccccacca 180  
 aaaataagct accatatagc ttataagtct caaatttttg ccttttacta aaatgtgatt 240  
 gtttctgttc attgtgtatg cttcatcacc tatattaggc aaattccatt ttttcccttg 300  
 t 301

<210> 287  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<400> 287  
 tacagatctg ggaactaaat attaaaaatg agtgtggctg gatatatgga gaatgttggg 60  
 cccagaagga acgtagagat cagatattac aacagctttg tttgagggg tagaaatatg 120  
 aaatgatttg gttatgaacg cacagtttag gcagcagggc cagaatcctg accctctgcc 180  
 ccgtggttat ctcctcccca gcttggctgc ctcagtgtat cacagtattc cattttgttt 240  
 gttgcatgtc ttgtgaagcc atcaagattt tctcgtctgt tttcctctca ttggtaatgc 300  
 t 301

<210> 288  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<400> 288  
 gtacacctaa ctgcaaggac agctgaggaa tgtaatgggc agccgctttt aaagaagtag 60  
 agtcaatagg aagacaaatt ccagttccag ctcagtctgg gtatctgcaa agctgcaaaa 120  
 gatcttttaa gacaatttca agagaatatt tccttaaagt tggcaatttg gagatcatac 180  
 aaaagcatct gcttttgtga tttaatttag ctcactctggc cactggaaga atccaaacag 240  
 tctgccttaa ttttgatga atgcatgatg gaaattcaat aatttagaaa gttaaaaaaa 300  
 a 301

<210> 289  
 <211> 301  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(301)  
 <223> n = A,T,C or G

<400> 289



```

ggtagactgt ttccatgtta tgtttctaca cattgctacc tcagtgtccc tggaaactta      60
gcttttgatg tctccaagta gtccaccttc atttaactct ttgaaactgt atcatctttg      120
ccaagtaaga gtggtggcct atttcagctg ctttgacaaa atgactggct cctgacttaa      180
cgttctataa atgaatgtgc tgaagcaaag tgcccatggg gccggcgaaan aagagaaaga      240
tgtgttttgt tttggactct ctgtgggtccc ttccaatgct gtgggtttcc aaccagnnga      300
a                                                                              301

```

```

<210> 290
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (301)
<223> n = A,T,C or G

```

```

<400> 290
acactgagct ottcttgata aatatacaga atgcttggca tatacaagat tctatactac      60
tgactgatct gttcatttct ctcacagctc ttaccccaa aagcttttcc accctaagtg      120
ttctgacctc ctttttctaata cacagtaggg atagaggcag anccacctac aatgaacatg      180
gagttctatc aagaggcaga aacagcacag aatcccagtt ttaccattcg ctagcagtgc      240
tgccttgaac aaaaacattt ctccatgtct ctttttcttc atgcctcaag taacagtgcg      300
a                                                                              301

```

```

<210> 291
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 291
caggtaccaa tttcttctat cctagaaaca tttcatttta tgttgttgaa acataacaac      60
tatatcagct agattttttt tctatgcttt acctgctatg gaaaatttga cacattctgc      120
tttactcttt tgtttatagg tgaatcacia aatgtatttt tatgtattct gtagttcaat      180
agccatggct gtttacttca ttttaatttt ttagcataaa gacattatga aaaggcctaa      240
acatgagctt cacttcccca ctaactaatt agcatctggt atttcttaac cgtaatgcct      300
a                                                                              301

```

```

<210> 292
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1) ... (301)
<223> n = A,T,C or G

```

```

<400> 292
accttttagt agtaatgtct aataataaat aagaaatcaa ttttataagg tccatatagc      60
tgtattaaat aatttttaag tttaaaagat aaaataccat catttttaaat gttggtattc      120

```



```
<210> 293
<211> 301
<212> DNA
<213> Homo sapien
```

```
<210> 294
<211> 301
<212> DNA
<213> Homo sapien
```

```
<210> 295
<211> 305
<212> DNA
<213> Homo sapien
```

```
<210> 296
<211> 301
```



<400> 296

<210> 297

<220>

<400> 297

<210> 298

 $\langle 220 \rangle$ 

<400> 298

<210> 299

```
<211> 301
<212> DNA
<213> Homo sapien
```



gttttgagac	ggagttttcac	tcttgttgcc	cagactggac	tgcaatggca	gggtctctgc	60
tactgcacc	ctctgcctcc	caggttcgag	caattctcct	gcctcagcct	cccaggtagc	120
tgggattgca	ggctcacgcc	accataccca	gctaattttt	ttgtattttt	agtagagacg	180
gagtttctgc	atgttggcca	gctgggtctca	aactcctgac	ctcaagcgac	ctgcctgcct	240
cggcctccca	aagtgtctgga	attataggca	tgagtcaaca	cgccccagcct	aaagatatatt	300
t						301

<213> Homo sapien

attcagtttt	atttgetgcc	ccagtatctg	taaccaggag	tgccacaaaa	tcttgccaga	60
tatgtcccac	accactggg	aaaggctccc	acctggctac	ttcctctatc	agctgggtca	120
gctgcattcc	acaaggttct	cagcctaattg	agtttcacta	cctgccagtc	tcaaaaactta	180
gtaaagcaag	accatgacat	tccccacgg	aatcagagt	ttgccccacc	gtcttgttac	240
tataaagcct	gcctctaaca	gtccttgctt	cttcacacca	atcccgagcg	catcccccat	300
g						301

<213> Homo sapien

ttaaattttt	gagaggataa	aaaggacaaa	taatctagaa	atgtgtcttc	ttcagttctgc	60
agaggacccc	aggtctccaa	gcaaccacat	ggtcaagggc	atgaataatt	aaaagttgggt	120
gggaactcac	aaagaccctc	agagctgaga	caccacaaac	agtgggagct	cacaaagacc	180
ctcagagctg	agacacccac	aacagtggga	gctcaciaag	accctcagag	ctgagacacc	240
cacaacagca	cctcgttcag	ctgccacatg	tgtgaataag	gatgcaatgt	ccagaagtgt	300
t						301

<213> Homo sapien

aggtacacat	ttagcttg	gtaaatgact	cacaaaactg	attttaaaat	caagttaatg	60
tgaattttga	aaattactac	ttaatcctaa	ttcacataa	caatggcatt	aaggtttgac	120
ttgagttggt	tcttagtatt	atztatggta	aataggctct	taccacttgc	aaataactgg	180
ccacatcatt	aatgactgac	ttcccagtaa	ggctctctaa	ggggtaaagta	ggaggatcca	240
caggatttga	gatgctaagg	ccccagagat	cgtttgatcc	aaccctctta	ttttcagagg	300
g						301

<213> Homo sapien



```
<210> 307
<211> 637
```







```
<210> 310
<211> 539
<212> DNA
<213> Homo sapien
```

```
<210> 311
<211> 526
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1) ... (526)
<223> n = A,T,C or G
```

```
<210> 312
<211> 500
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1) ... (500)
<223> n = A,T,C or G
```

<400> 312



```

cctctctctc cccaccccct gactctagag aactggggtt tctcccagta ctccagcaat      60
tcattttctga aagcagttga gccactttat tccaaagtac actgcagatg ttcaaactct      120
ccattttctct ttcccttcca cctgccagtt ttgctgactc tcaacttgtc atgagtgtaa      180
gcattaagga cattatgctt cttcgattct gaagacaggc cctgctcatg gatgactctg      240
gcttcttagg aaaatatttt tcttccaaaa tcagtaggaa atctaaactt atccccctct      300
tgcagatgtc tagcagcttc agacatttgg ttaagaacct atgggaaaaa aaaaaatcct      360
tgctaattgt gtttcccttg taaaccanga ttcttatttg nctggtatag aatatcagct      420
ctgaacgtgt ggtaaagatt tttgtgtttg aatataggag aaatcagttt gctgaaaagt      480
tagtcttaat tatctattgg                                     500

```

```

<210> 313
<211> 718
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(718)
<223> n = A,T,C or G

```

```

<400> 313
ggagatttgt gtggtttgca gccgaggag accaggaaga tctgcatggt gggaaggacc      60
tgatgataca gaggtgagaa ataagaaagg ctgctgactt taccatctga ggccacacat      120
ctgctgaaat ggagataatt aacatcacta gaaacagcaa gatgacaata taatgtctaa      180
gtagtacat gtttttgcac atttccagcc cttttaaata tccacacaca caggaagcac      240
aaaaggaagc acagagatcc ctgggagaaa tgcccggccg ccatcttggg tcatcgatga      300
gctctgccct gtgctgntc ccgcttgtga gggaaggaca ttagaaaatg aattgatgtg      360
ttccttaaag gatggcagga aaacagatcc tgttgtggat atttatttga acgggattac      420
agatttgaaa tgaagtcaca aagtgagcat taccaatgag aggaaaacag acgagaaaat      480
cttgatgggt cacaagacat gcaacaaaca aaatggaata ctgtgatgac acgagcagcc      540
aactggggag gagataccac ggggcagagg tcaggattct ggccctgctg cctaactgtg      600
cgttatacca atcatttcta tttctaccct caaacaagct gtngaatatc tgacttaacgg      660
ttcttntggc ccacatttct atnatccacc cctccttttt aannttantc caaantgt      718

```

```

<210> 314
<211> 358
<212> DNA
<213> Homo sapien

```

```

<400> 314
gtttattttac attacagaaa aaacatcaag acaatgtata ctatttcaaa tatatccata      60
cataatcaaa tatagctgta gtacatgttt tcattggtgt agattaccac aaatgcaagg      120
caacatgtgt agatctcttg tcttattctt ttgtctataa tactgtattg tgtagtccaa      180
gctctcggtg gtccagccac tgtgaaacat gtcctcttta gattaacctc gtggacgctc      240
ttgttgatt gctgaactgt agtgcctgt attttgcttc tgtctgtgaa ttctgttgct      300
tctggggcat ttccttgtga tgcagaggac caccacacag atgacagcaa tctgaatt      358

```

```

<210> 315
<211> 341
<212> DNA
<213> Homo sapien

```



&lt;400&gt; 315

taccacctcc	ccgctggcac	tgatgagccg	catcaccatg	gtcaccagca	ccatgaaggc	60
ataggtgatg	atgaggacat	ggaatgggcc	cccaaggatg	gtctgtccaa	agaagcgagt	120
gacccccatt	ctgaagatgt	ctggaacctc	taccagcagg	atgatgatag	ccccaatgac	180
agtcaccagc	tccccgacca	gccggatatc	gtccttaggg	gtcatgtagg	cttcctgaag	240
tagcttctgc	tgtaagaggg	tgttgtcccg	ggggctcgtg	cggttattgg	tcttgggctt	300
gagggggcgg	tagatgcagc	acatggtgaa	gcagatgatg	t		341

&lt;210&gt; 316

&lt;211&gt; 151

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 316

agactgggca	agactcttac	gccccacact	gcaatttggt	cttggtgccc	tatccattta	60
tgtgggcctt	tctcgagttt	ctgattataa	acaccactgg	agcgatgtgt	tgactggact	120
cattcaggga	gctctgggtg	caatattagt	t			151

&lt;210&gt; 317

&lt;211&gt; 151

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 317

agaactagt	gataccta	aaataacctga	aacatatatt	ggcattttatc	aatggctcaa	60
atcttcattt	atctctggcc	ttaaccctgg	ctcctgaggg	tgccggccagc	agatcccagg	120
ccagggtct	gttcttgcca	cacctgcttg	a			151

&lt;210&gt; 318

&lt;211&gt; 151

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 318

actggtggga	ggcgcgtgtt	agttggctgt	tttcagaggg	gtctttcggga	gggacctcct	60
gctgcaggct	ggagtgtctt	tattcctggc	gggagaccgc	acattccact	gctgaggctg	120
tgggggcggt	ttatcaggca	gtgataaaca	t			151

&lt;210&gt; 319

&lt;211&gt; 151

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 319

aactagtgga	tccagagcta	taggtacagt	gtgatctcag	ctttgcaaac	acattttcta	60
catagatagt	actaggtatt	aatagatatg	taaagaaaga	aatcacacca	ttaataatgg	120
taagattggg	tttatgtgat	tttagtgggg	a			151

&lt;210&gt; 320

&lt;211&gt; 150



<400> 320

<400> 321

 $\langle 220 \rangle$ 

<400> 322

 $\langle 220 \rangle$ 

<400> 323

<210>	324
<211>	461
<212>	DNA



<213> Homo sapien

<220>

<221> misc\_feature

<222> (1) ... (461)

<223> n = A,T,C or G

<400> 324

acctgtgtgg	aatttcagct	ttcctcatgc	aaaaggattt	tgtatccccg	gcctacttga	60
agaagtggtc	agctaaagga	atccagggtg	ttggttggac	tgtaataacc	tttgatgaaa	120
agagttacta	cgaatcccat	cttgggtcca	gctatatcac	tgacagcatg	gtagaagact	180
gcgaacctca	cttctagact	ttcacgggtg	gacgaaacgg	gttcagaaac	tgccaggggc	240
ctcatacagg	gatatacaaa	taccctttgt	gctaccagg	ccctggggaa	tcagggtgact	300
cacacaaatg	caatagttgg	tactgcatt	tttacctgaa	ccaaagctaa	acccgggtgtt	360
gccaccatgc	accatggcat	gccagagttc	aacactgttg	ctcttgaaaa	ttgggtctga	420
aaaaacgcac	aagagcccct	gcctgcct	agctgangca	c		461

<210> 325

<211> 400

<212> DNA

<213> Homo sapien

<400> 325

acactgtttc	catgttatgt	ttctacacat	tgctacctca	gtgctcctgg	aaacttagct	60
tttgatgtct	ccaagtagtc	caccttcatt	taactctttg	aaactgtatc	atctttgcc	120
agtaagagtg	gtggcctatt	tcagctgctt	tgacaaaatg	actggctcct	gacttaacgt	180
tctataaatg	aatgtgctga	agcaaagtgc	ccatgggtggc	ggcgaagaag	agaaagatgt	240
gttttgtttt	ggactctctg	tggtcccttc	caatgctgtg	ggtttccaac	caggggaagg	300
gtcccttttg	cattgccaaag	tgccataacc	atgagcacta	cgctaccatg	gttctgcctc	360
ctggccaagc	aggctggttt	gcaagaatga	aatgaatgat			400

<210> 326

<211> 1215

<212> DNA

<213> Homo sapien

<400> 326

ggaggactgc	agcccgcact	cgcagccctg	gcaggcggca	ctggatcatgg	aaaacgaatt	60
gttctgctcg	ggcgtcctgg	tgcatccgca	gtgggtgctg	tcagccgcac	actgtttcca	120
gaactcctac	accatcgggc	tgggcctgca	cagtcttgag	gccgaccaag	agccagggag	180
ccagatgggtg	gaggccagcc	tctccgtacg	gcacccagag	tacaacagac	ccttgctcgc	240
taacgacctc	atgctcatca	agttggacga	atccgtgtcc	gagtctgaca	ccatccggag	300
catcagcatt	gcttcgcagt	gccctaccgc	ggggaaactct	tgctcgtttt	ctggctgggg	360
tctgctggcg	aacggcagaa	tgctaccgt	gctgcagtgc	gtgaacgtgt	cgggtggtgtc	420
tgaggaggtc	tgagtaagc	tctatgaccc	gctgtaccac	cccagcatgt	tctgcgccgg	480
cggaggggcaa	gaccagaagg	actcctgcaa	cggtgactct	ggggggcccc	tgatctgcaa	540
cgggtacttg	cagggccttg	tgtctttcgg	aaaagccccg	tgtggccaag	ttggcgtgcc	600
aggtgtctac	accaacctct	gcaaattcac	tgagtggata	gagaaaaccg	tccaggccag	660
ttaactctgg	ggactgggaa	cccatgaaat	tgacccccaa	atacatcctg	cgggaaggaat	720
tcaggaatat	ctgttcccag	cccctcctcc	ctcaggcccc	ggagtccagg	ccccagccc	780
ctcctccctc	aaaccaaggg	tacagatccc	cagccctcc	tccctcagac	ccaggagtcc	840



```

agacccccca gcccctcctc cctcagaccc aggagtccag cccctcctcc ctcagaccca      900
ggagtccaga cccccagcc cctcctccct cagaccaggg ggtccaggcc cccaaccct      960
cctccctcag actcagaggt ccaagcccc aaccctcct tccccagacc cagagggtcca    1020
ggtcccagcc cctcctccct cagaccagc ggtccaatgc cacctagact ctccctgtac    1080
acagtgcgcc cttgtggcac gttgacccaa ccttaccagt tggtttttca ttttttgtcc    1140
ctttccccta gatccagaaa taaagtctaa gagaagcgca aaaaaaaaaa aaaaaaaaaa    1200
aaaaaaaaaa aaaaaa                                     1215

```

```

<210> 327
<211> 220
<212> PRT
<213> Homo sapien

```

```

<400> 327
Glu Asp Cys Ser Pro His Ser Gln Pro Trp Gln Ala Ala Leu Val Met
1      5      10      15
Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln Trp Val
20     25     30
Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu Gly
35     40     45
Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met Val Glu
50     55     60
Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu Leu Ala
65     70     75     80
Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser Asp
85     90     95
Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly Asn
100    105    110
Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg Met Pro
115    120    125
Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu Glu Val Cys
130    135    140
Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys Ala Gly
145    150    155    160
Gly Gly Gln Asp Gln Lys Asp Ser Cys Asn Gly Asp Ser Gly Gly Pro
165    170    175
Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly Lys Ala
180    185    190
Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn Leu Cys Lys
195    200    205
Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
210    215    220

```

```

<210> 328
<211> 234
<212> DNA
<213> Homo sapien

```

```

<400> 328
cgctcgtctc tggtagctgc agccaaatca taaacggcga ggactgcagc ccgcactcgc      60
agccctggca ggcggcactg gtcattgaaa acgaattggt ctgctcgggc gtctctggtgc    120

```



<400>	332						
tggtgccgct	gcagccggca	gagatggttg	agctcatgtt	cccgtgttg	ctcctccttc		60
tgcccttcct	tctgtatatg	gctgcgcccc	aaatcaggaa	aatgctgtcc	agtgggggtgt		120
gtacatcaac	tgttcagctt	cctgggaaag	tagttgtggt	cacaggagct	aatacaggta		180
tcgggaagga	gacagccaaa	gagctggctc	agagaggagc	tcgagtatat	ttagcttgcc		240
gggatgtgga	aaagggggaa	ttggtggcca	aagagatcca	gaccacgaca	gggaaccagc		300



```
<210> 333
<211> 3030
<212> DNA
<213> Homo sapien
```

<400> 333							
gcaggcgact	tgcgagctgg	gagcgattta	aaacgctttg	gattcccccg	gcctgggttg		60
ggagagcgag	ctgggtgcc	cctagattcc	cgcgcccg	acctcatgag	ccgaccctcg		120
gctccatgga	gcccgcaat	tatgccacct	tggatggagc	caaggatata	gaaggcttgc		180
tgggagcggg	agggggcg	aatctggctg	ccactccc	tctgaccagc	caccagcgg		240
cgcctacgct	gatgcctgct	gtcaactatg	cccccttggg	tctgccaggc	tggcgaggc		300
cgccaaagca	atgccacca	tgccctgggg	tgccccagg	gacgtcccca	gctcccgctg		360
cttatggtta	ctttggaggc	gggtactact	cctgccgaqt	gtcccgaagc	tcqctqaaac		420



cctgtgcccc	ggcagccacc	ctggccgcgt	accccgcgga	gactcccacg	gccggggaag	480
agtaccccag	ycgccccact	gagtttgcc	tctatccggg	atatccggga	acctaccagc	540
ctatggccag	ttacctggac	gtgtctgtgg	tgcagactct	gggtgtccct	ggagaaccgc	600
gacatgactc	cctgttgcc	gtggacagtt	accagtcttg	ggctctcgt	ggtggtgga	660
acagccagat	gtgttgccag	ggagaacaga	acccaccagg	tcccttttgg	aaggcagcat	720
ttgcagactc	cagcgggcag	cacctctctg	acgcctgcgc	ctttcgtcgc	ggccgcaaga	780
aacgcattcc	gtacagcaag	gggcagttgc	gggagctgga	gccggagtat	gcggctaaca	840
agttcatcac	caaggacaag	aggcgcaaga	tctcggcagc	caccagcctc	tcggagcgcc	900
agattacat	ctggtttcag	aaccgccggg	tcaaagagaa	gaaggttctc	gccaaagtga	960
agaacagcgc	taccctttaa	gagatctcct	tgcttggttg	ggaggagcga	aagtgggggt	1020
gtcctgggga	gaccaggaac	ctgccaaagc	caggctgggg	ccaaggactc	tgctgagagg	1080
cccctagaga	caacaccctt	cccaggccac	tggtgtctgg	actgttcctc	aggagcggcc	1140
tggttaccca	gtatgtgcag	ggagacggaa	ccccatgtga	cagccactc	caccagggtt	1200
cccaaagaac	ctggcccagt	cataatcatt	catctgaca	gtggcaataa	tcacgataac	1260
cagtactagc	tgccatgac	gttagcctca	tattttctat	ctagagctct	gtagagcact	1320
ttagaaaccg	ctttcatgaa	ttgagcta	tatgaataaa	tttggaaagg	gatccctttg	1380
cagggaagct	ttctctcaga	cccccttcca	ttacacctct	cacctggta	acagcaggaa	1440
gactgaggag	aggggaacgg	gcagattcgt	tgtgtggctg	tgatgtccgt	ttagcatttt	1500
tctcagctga	cagctgggta	ggtggacaat	tgtagaggct	gtctcttcc	ccctccttgt	1560
ccaccccata	gggtgtaccc	actggtcttg	gaagcaccca	tccttaatac	gatgatTTTT	1620
ctgtcgtgtg	aaaatgaagc	cagcaggctg	cccctagtca	gtccttccct	ccagagaaaa	1680
agagatttga	gaaagtgcct	gggtaattca	ccattaattt	cctcccccaa	actctctgag	1740
tcttccctta	atatttcttg	tggttctgac	caaagcaggt	catggtttgt	tgagcatttg	1800
ggatcccagt	gaagtagatg	tttgtagcct	tgcatactta	gcccttccca	ggcacaaaac	1860
gagtggcaga	gtggtgccaa	ccctgttttc	ccagtccacg	tagacagatt	cacagtgcgg	1920
aattctggaa	gctggagaca	gacgggctct	ttgcagagcc	gggactctga	gagggacatg	1980
agggcctctg	cctctgtgtt	cattctctga	tgtcctgtac	ctgggctcag	tgcccgggtg	2040
gactcatctc	ctggccgcgc	agcaaagcca	gcgggttcgt	gctggctcct	cctgcacctt	2100
aggtctgggg	tggggggcct	gccggcgcat	tctccacgat	tgagcgaca	ggcctgaagt	2160
ctggacaacc	cgcagaaccg	aagctccgag	cagcgggtcg	gtggcgagta	gtggggctcg	2220
tggcgagcag	ttggtggtgg	gccgcggccg	ccactacctc	gaggacattt	ccctcccgga	2280
gccagctctc	ctagaaaccc	cgcggcgggc	gccgcagcca	agtgtttatg	gcccgcgggtc	2340
gggtgggatc	ctagccctgt	ctcctctcct	gggaaggagt	gaggggtggga	cgtgacttag	2400
acacctacaa	atctattttac	caaagaggag	cccgggactg	agggaaaagg	ccaaagagtg	2460
tgagtgcattg	cggactgggg	gttcagggga	agaggacgag	gaggaggaag	atgaggtcga	2520
tttcttgatt	taaaaaatcg	tccaagcccc	gtgtccagc	ttaaggctct	cggttacatg	2580
cgcgcctcag	agcaggtcac	tttctgcctt	ccacgtcctc	cttcaaggaa	gccccatgtg	2640
ggtagctttc	aatatcgag	gttcttactc	ctctgcctct	ataagctcaa	acccaccaac	2700
gatcgggcaa	gtaaaccccc	tcctctcgcc	acttcggaac	tggcgagagt	tcagcgcaga	2760
tgggcctgtg	gggagggggc	aagatagatg	agggggagcg	gcatggtgcg	gggtgacccc	2820
ttggagagag	gaaaaaggcc	acaagagggg	ctgccaccgc	cactaacgga	gatggccctg	2880
gtagagacct	ttgggggtct	ggaacctctg	gactccccat	gctctaactc	ccacactctg	2940
ctatcagaaa	cttaaaacttg	aggattttct	ctgtttttca	ctcgcaataa	aytcagagca	3000
aacaaaaaaaa	aaaaaaaaaa	aaaactcgag				3030

&lt;210&gt; 334

&lt;211&gt; 2417

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 334



ggcgcccgct ctagagctag tgggatcccc cgggctgcac gaattcggca cgagtgaatt 60  
 ggagttttac ctgtattgtt ttaatttcaa caagcctgag gactagccac aaatgtaccc 120  
 agttttacaaa tgaggaaaca ggtgcaaaaa ggttgttacc tgtcaaaggc cgtatgtggc 180  
 agagccaaga tttgagccca gttatgtctg atgaacttag cctatgctct ttaaacttct 240  
 gaatgctgac cattgaggat atctaaactt agatcaattg cattttccct ccaagactat 300  
 ttacttatca atacaataat accaccttta ccaatctatt gttttgatac gagactcaaa 360  
 tatgccagat atatgtaaaa gcaacctaca agctctctaa tcatgctcac ctaaaagatt 420  
 cccgggatct aataggctca aagaaacttc ttctagaaat ataaaagaga aaattggatt 480  
 atgcaaaaat tcattattaa tttttttcat ccatccttta attcagcaaa catttatctg 540  
 ttgttgactt tatgcagtat ggccttttaa ggattggggg acagggtgaag aacgggggtgc 600  
 cagaatgcat cctcctacta atgaggctag tacacatttg cattttaaaa tgccctgtcc 660  
 agctgggcat ggtggatcat gcctgtaatc tcaacattgg aaggccaagg caggaggatt 720  
 gcttcagccc aggagttcaa gaccagcctg ggcaacatag aaagacccca tctctcaatc 780  
 aatcaatcaa tgccctgtct ttgaaaataa aactccttaa gaaaggttta atgggcaggg 840  
 tgtggtagct catgcctata atacagcact ttgggaggct gaggcaggag gatcacttta 900  
 gcccagaagt tcaagaccag cctgggcaac aagtgcacc tcatctcaat tttttaataa 960  
 aatgaataca tacataagga aagataaaaa gaaaagttaa atgaaagaat acagtataaa 1020  
 acaaatctct tggacctaaa agtatttttg ttcaagccaa atattgtgaa tcacctctct 1080  
 gtgttgagga tacagaatat ctaagccag gaaactgagc agaaagtca tgtactaact 1140  
 aatcaacccg aggcaaggca aaaatgagac taactaatca atccgaggca aggggcaaat 1200  
 tagacggaac ctgactctgg tctattaagc gacaactttc cctctgttgt atttttcttt 1260  
 tattcaatgt aaaaggataa aaactctcta aaactaaaaa caatgtttgt caggagttac 1320  
 aaaccatgac caactaatta tggggaatca taaaatatga ctgtatgaga tcttgatggg 1380  
 ttacaaagtg taccactgt taatcacttt aaacattaat gaacttaaaa atgaatttac 1440  
 ggagattgga atgtttcttt cctgttgtat tagttggctc aggtgccat aacaaaatac 1500  
 cacagactgg gaggcttaag taacagaaat tcatttctca cagttctggg ggctggaagt 1560  
 ccacgatcaa ggtgcaggaa aggcaggctt cattctgagg cccctctctt ggctcacatg 1620  
 tggccacctt cccactgctt gctcacatga cctctttgtg ctctggaaa gaggggtgtg 1680  
 gggacagagg gaaagagaag gagagggaac tctctggtgt ctctgtttc aaggacccta 1740  
 acctgggcca ctttggccca ggcactgtgg ggtggggggg tgtggctgct ctgctctgag 1800  
 tggccaagat aaagcaacag aaaaatgtcc aaagctgtgc agcaaagaca agccaccgaa 1860  
 cagggatctg ctcatcagtg tggggacctc caagtcggcc accctggagg caagccccc 1920  
 cagagcccat gcaagggtggc agcagcagaa gaagggaatt gtccctgtcc ttggcacatt 1980  
 cctcaccgac ctggtgatgc tggacactgc gatgaatggg aatgtggatg agaatatgat 2040  
 ggactcccag aaaaggagac ccagctgtc aggtggctgc aaatcattac agccttcac 2100  
 ctggggagga actggggggc tggttctggg tcagagagca gccagtgag ggtgagagct 2160  
 acagcctgtc ctgccagctg gatccccagt cccggtcaac cagtaatcaa ggtgagcag 2220  
 atcaggcttc ccggagctgg tcttggaag ccagccctgg ggtgagttgg ctctgctgt 2280  
 ggtactgaga caatattgtc ataaattcaa tgcgcccttg tatccctttt tcttttttat 2340  
 ctgtctacat ctataatcac tatgcatact agtctttgtt agtgtttcta ttcmaactaa 2400  
 tagagatatg ttatact 2417

<210> 335

<211> 2984

<212> DNA

<213> Homo sapien

<400> 335

atccctcctt cccactctc ctttccagaa ggcaattggg gtcttatctg ttggactctg 60  
 aaaacacttc aggcgccctt ccaaggcttc cccaaacccc taagcagccg cagaagcgt 120  
 cccgagctgc cttctccac actcaggtga tcgagttgga gaggaagttc agccatcaga 180







<213> Homo sapien

<400> 336

```

Pro Ser Phe Pro Thr Leu Leu Ser Arg Arg His Leu Gly Ser Tyr Leu
 1          5          10          15
Leu Asp Ser Glu Asn Thr Ser Gly Ala Leu Pro Arg Leu Pro Gln Thr
          20          25          30
Pro Lys Gln Pro Gln Lys Arg Ser Arg Ala Ala Phe Ser His Thr Gln
          35          40          45
Val Ile Glu Leu Glu Arg Lys Phe Ser His Gln Lys Tyr Leu Ser Ala
 50          55          60
Pro Glu Arg Ala His Leu Ala Lys Asn Leu Lys Leu Thr Glu Thr Gln
65          70          75          80
Val Lys Ile Trp Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln
          85          90          95
Leu Ser Ser Glu Leu Gly Asp Leu Glu Lys His Ser Ser Leu Pro Ala
          100          105          110
Leu Lys Glu Glu Ala Phe Ser Arg Ala Ser Leu Val Ser Val Tyr Asn
          115          120          125
Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr Cys Val Gly Ser Trp Ser Pro
          130          135          140
Ala Phe Trp
145

```

<210> 337

<211> 9

<212> PRT

<213> Homo sapien

<400> 337

```

Ala Leu Thr Gly Phe Thr Phe Ser Ala
 1          5

```

<210> 338

<211> 9

<212> PRT

<213> Homo sapien

<400> 338

```

Leu Leu Ala Asn Asp Leu Met Leu Ile
 1          5

```

<210> 339

<211> 318

<212> PRT

<213> Homo sapien

<400> 339

```

Met Val Glu Leu Met Phe Pro Leu Leu Leu Leu Leu Pro Phe Leu
 1          5          10          15

```



Leu Tyr Met Ala Ala Pro Gln Ile Arg Lys Met Leu Ser Ser Gly Val  
                   20                  25                  30  
 Cys Thr Ser Thr Val Gln Leu Pro Gly Lys Val Val Val Val Thr Gly  
                   35                  40                  45  
 Ala Asn Thr Gly Ile Gly Lys Glu Thr Ala Lys Glu Leu Ala Gln Arg  
                   50                  55                  60  
 Gly Ala Arg Val Tyr Leu Ala Cys Arg Asp Val Glu Lys Gly Glu Leu  
 65                  70                  75                  80  
 Val Ala Lys Glu Ile Gln Thr Thr Thr Gly Asn Gln Gln Val Leu Val  
                   85                  90                  95  
 Arg Lys Leu Asp Leu Ser Asp Thr Lys Ser Ile Arg Ala Phe Ala Lys  
                   100                  105                  110  
 Gly Phe Leu Ala Glu Glu Lys His Leu His Val Leu Ile Asn Asn Ala  
                   115                  120                  125  
 Gly Val Met Met Cys Pro Tyr Ser Lys Thr Ala Asp Gly Phe Glu Met  
 130                  135                  140  
 His Ile Gly Val Asn His Leu Gly His Phe Leu Leu Thr His Leu Leu  
 145                  150                  155                  160  
 Leu Glu Lys Leu Lys Glu Ser Ala Pro Ser Arg Ile Val Asn Val Ser  
                   165                  170                  175  
 Ser Leu Ala His His Leu Gly Arg Ile His Phe His Asn Leu Gln Gly  
                   180                  185                  190  
 Glu Lys Phe Tyr Asn Ala Gly Leu Ala Tyr Cys His Ser Lys Leu Ala  
                   195                  200                  205  
 Asn Ile Leu Phe Thr Gln Glu Leu Ala Arg Arg Leu Lys Gly Ser Gly  
 210                  215                  220  
 Val Thr Thr Tyr Ser Val His Pro Gly Thr Val Gln Ser Glu Leu Val  
 225                  230                  235                  240  
 Arg His Ser Ser Phe Met Arg Trp Met Trp Trp Leu Phe Ser Phe Phe  
                   245                  250                  255  
 Ile Lys Thr Pro Gln Gln Gly Ala Gln Thr Ser Leu His Cys Ala Leu  
                   260                  265                  270  
 Thr Glu Gly Leu Glu Ile Leu Ser Gly Asn His Phe Ser Asp Cys His  
                   275                  280                  285  
 Val Ala Trp Val Ser Ala Gln Ala Arg Asn Glu Thr Ile Ala Arg Arg  
                   290                  295                  300  
 Leu Trp Asp Val Ser Cys Asp Leu Leu Gly Leu Pro Ile Asp  
 305                  310                  315

<210> 340

<211> 483

<212> DNA

<213> Homo sapien

<400> 340

gccgaggtct gccttcacac ggaggacacg agactgcttc ctcaagggct cctgacctgcc 60  
 tggacactgg tgggaggcgc tgtttagttg gctgttttca gagggggtctt tcggagggac 120  
 ctctgctgc aggctggagt gtctttattc ctggcgggag accgcacatt ccaactgctga 180  
 ggttggtggg gcggtttatc aggcagtgat aaacataaga tgtcatttcc ttgactccgg 240  
 ctttcaattt tctctttggc tgacgacgga gtccgtggtg tcccgatgta actgaccct 300  
 gctccaaacg tgacatcact gatgctcttc tcgggggtgc tgatggcccg cttggtcacg 360



```

tgetcaatct cgccattcga ctcttgctcc aaactgtatg aagacacctg actgcacgtt 420
ttttctgggc ttccagaatt taaagtgaag ggcagcactc ctaagctccg actccgatgc 480
ctg 483

```

```

<210> 341
<211> 344
<212> DNA
<213> Homo sapien

```

```

<400> 341
ctgtgctga gtcacagatt tcattataaa tagcctccct aaggaaaata cactgaatgc 60
tatttttact aaccattcta tttttataga aatagctgag agtttctaaa ccaactctct 120
gctgccttac aagtattaaa tattttactt ctttccataa agagtagctc aaaatatgca 180
attaatttaa taatttctga tgatggtttt atctgcagta atatgtatat catctattag 240
aatttactta atgaaaaact gaagagaaca aaatttgtaa ccactagcac ttaagtactc 300
ctgattctta acattgtctt taatgaccac aagacaacca acag 344

```

```

<210> 342
<211> 592
<212> DNA
<213> Homo sapien

```

```

<400> 342
acagcaaaaa agaaactgag aagcccaaty tgctttcttg ttaacatcca cttatccaac 60
caatgtggaa acttcttata ctgggttcca ttatgaagtt ggacaattgc tgctatcaca 120
cctggcaggt aaaccaatgc caagagagtg atggaaacca ttggcaagac tttgttgatg 180
accaggattg gaattttata aaaatattgt tgatgggaag ttgctaaagg gtgaattact 240
tccttcagaa gagtgtaaaag aaaagtcaga gatgctataa tagcagctat ttttaattggc 300
aagtgccact gtggaaagag ttctgtgtgt tgctgaagtt ctgaagggca gtcaaattca 360
tcagcatggg ctgtttggtg caaatgcaaa agcacaggctc ttttttagcat gctggtctct 420
cccgtgtcct tatgcaaata atcgtcttct tctaaatttc tcctaggctt cattttccaa 480
agttcttctt ggtttgtgat gtcttttctg ctttccatta attctataaa atagtatggc 540
ttcagccacc cactcttcgc cttagcttga ccgtgagtct cggctgccgc tg 592

```

```

<210> 343
<211> 382
<212> DNA
<213> Homo sapien

```

```

<400> 343
ttcttgacct cctcctcctt caagctcaaa caccacctcc cttattcagg accggcactt 60
cttaatgttt gtggctttct ctccagcctc tcttaggagg ggtaatggtg gagttggcat 120
cttgtaactc tcctttctcc tttctctgcc cgcttttccc atcctgctgt 180
agacttcttg attgtcagtc tgtgtcacat ccagtgattg ttttggtttc tgttcccttt 240
ctgactgccc aaggggctca gaacccagc aatcccttcc tttcactacc ttcttttttg 300
ggggtagttg gaagggactg aaattgtggg ggggaaggtag gaggcacatc aataaaggag 360
aaaccaccaa gctgaaaaaa aa 382

```

```

<210> 344
<211> 536
<212> DNA

```



<223> n = A, T, C or G



&lt;400&gt; 347

acacacataa	tattataaaa	tgccatctaa	ttggaaggag	ctttctatca	ttgcaagtca	60
taaatataac	ttttaaaana	ntactancag	cttttaccta	ngctcctaaa	tgcttgtaaa	120
tctgagactg	actggacceca	cccagacceca	gggcaaagat	acatgttacc	atatcatctt	180
tataaagaat	ttttttttgt	c				201

&lt;210&gt; 348

&lt;211&gt; 251

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 348

ctgttaataca	caacattttgt	gcataccttg	tgccaagtga	gaaaatgttc	taaaatcaca	60
agagagaaca	gtgccagaat	gaaactgacc	ctaagtccca	ggtgcccctg	ggcaggcaga	120
aggagacact	cccagcatgg	aggagggttt	atcttttcat	cctagggtcag	gtctacaatg	180
ggggaagggtt	ttattataga	actccaaca	gccacactca	ctcctgccac	ccacccgatg	240
gccctgcttc	c					251

&lt;210&gt; 349

&lt;211&gt; 251

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 349

taaaaatcaa	gccattttaat	tgtatctttg	aaggtaaaca	atatatggga	gctggatcac	60
aacccctgag	gatgccagag	ctatgggtcc	agaacatggt	gtggtattat	caacagagtt	120
cagaagggtc	tgaactctac	gtgttaccag	agaacataat	gcaattcatg	cattccactt	180
agcaattttg	taaaatacca	gaaacagacc	ccaagagtct	ttcaagatga	ggaaaattca	240
actcctgggtt	t					251

&lt;210&gt; 350

&lt;211&gt; 908

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 350

ctggacactt	tgcgagggct	tttgcctggc	gctgctgctg	cccgtcatgc	tactcatcgt	60
agcccgcccc	gtgaagctcg	ctgctttccc	tacctcctta	agtgactgcc	aaacgccccac	120
cggctggaat	tgctctggtt	atgatgacag	agaaaatgat	ctcttcctct	gtgacaccaa	180
cacctgtaaa	tttgatgggg	aatgtttaag	aattggagac	actgtgactt	gcgtctgtca	240
gttcaagtgc	aacaatgact	atgtgcctgt	gtgtggctcc	aatggggaga	gctaccagaa	300
tgagtgttac	ctgcgacagg	ctgcatgcaa	acagcagagt	gagatacttg	tggtgtcaga	360
aggatcatgt	gccacagtcc	atgaaggctc	tgagaaaact	agtcaaaagg	agacatccac	420
ctgtgatatt	tgccagtttg	gtgcagaatg	tgacgaagat	gccgaggatg	tctggtgtgt	480
gtgtaatat	gactgtttct	aaaccaactt	caatcccctc	tgcgcttctg	atgggaaatc	540
ttatgataat	gcatgccaaa	tcaaagaagc	atcgtgtcag	aaacaggaga	aaattgaagt	600
catgtctttg	ggtcgatgtc	aagataacac	aactacaact	actaagtctg	aagatgggca	660
ttatgcaaga	acagattatg	cagagaatgc	taacaaatta	gaagaaagtg	ccagagaaca	720
ccacatacct	tgtccggaac	attacaatgg	cttctgcatg	catgggaagt	gtgagcattc	780
tatcaatatg	caggagccat	cttgcagggt	tgatgctggt	tatactggac	aacactgtga	840



aaaaaaggac tacagtgttc tatacgttgt tcccggtcct gtacgatttc agtatgtctt 900  
aatcgacag 908

<210> 351  
<211> 472  
<212> DNA  
<213> Homo sapien

<400> 351  
ccagttatatt gcaagtggta agagcctatt taccataaat aataactaaga accaactcaa 60  
gtcaaacctt aatgccattg ttattgtgaa ttaggattaa gtagtaattt tcaaaattca 120  
cattaacttg atttttaaata cagwtttgyg agtcatttac cacaagctaa atgtgtacac 180  
tatgataaaa acaaccattg tattcctggt ttcttaaaaca gtccctaattt ctaacactgt 240  
atatatcctt cgacatcaat gaactttggt ttcttttact ccagtaataa agtaggcaca 300  
gatctgtcca caacaaactt gccctctcat gccttgcttc tcaccatgct ctgctccagg 360  
tcagccccct tttggcctgt ttgttttgtc aaaaacctaa tctgcttctt gctttttcttg 420  
gtaatatata tttagggaag atgttgcttt gcccacacac gaagcaaagt aa 472

<210> 352  
<211> 251  
<212> DNA  
<213> Homo sapien

<400> 352  
ctcaaagcta atctctcggg aatcaaacca gaaaagggca aggatcttag gcatgggtgga 60  
tgtggataag gccagggtcaa tggctgcaag catgcagaga aagaggtaca tcggagcgtg 120  
caggctgcgt tccgtcctta cgatgaagac cagcatgcag tttccaaaca ttgccactac 180  
atacatggaa aggaggggga agccaaccca gaaatgggct ttctctaata ctgggataacc 240  
aataagcaca a 251

<210> 353  
<211> 436  
<212> DNA  
<213> Homo sapien

<400> 353  
tttttttttt tttttttttt ttttttataa caatgcagtc atttatttat tgagtatgtg 60  
cacattatgg tattattact atactgatta tatttatcat gtgacttcta attaraaaat 120  
gtatccaaaa gcaaaacagc agatatacaa aattaaagag acagaagata gacattaaca 180  
gataaggcaa cttatacatt gacaatccaa atccaatata tttaaactt tgggaaatga 240  
gggggacaaa tggaagccar atcaaatttg tgtaaaacta ttcagtatgt ttcccttgct 300  
tcatgtctga raaggctctc ccttcaatgg ggaatgacaaa ctccaaatgc cacacaaatg 360  
ttaacagaat actagattca cactggaacg ggggtaaaga agaaattatt ttctataaaa 420  
gggctcctaa tgtagt 436

<210> 354  
<211> 854  
<212> DNA  
<213> Homo sapien

<400> 354







<400> 357

<210> 358

<212> DNA

<400> 358

<210> 359

<212> DNA

<400> 359

acagcattcc	aaaatataca	tctagagact	aarrgtaa	gctctatagt	gaagaagtaa	60
taattaaaaa	atgctactaa	tatagaaaat	ttataatcag	aaaaataaat	attcagggag	120
ctcaccagaa	gaataaagtg	ctctgccagt	tattaaagga	ttactgctgg	tgaattaaat	180
atggcattcc	ccaagggaaa	tagagagatt	cttctggatt	atgttcaata	tttatttcac	240
aggattaact	gttttaggaa	cagatataaa	gcttcgccac	ggaagagatg	gacaaagcac	300
aaagacaaca	tgatacctta	ggaagcaaca	ctaccctttc	aggcataaaa	tttgagaaa	360
tgcaacatta	tgcttcatga	ataatatgta	gaaagaaggt	ctgatgaaaa	tgacatcctt	420
aatgtaaagt	aactttataa	gaattctggg	tcaaataaaa	ttctttgaag	aaaacatcca	480
aatgtcattg	acttatcaaa	tactatcttg	gcataatacc	tatgaaggca	aaactaaaca	540
aacaaaaagc	tcacacaaaa	caaaaccatc	aacttatttt	gtattctata	acatacgaga	600
ctgtaaagat	gtgacagtgt					620



<210> 360  
 <211> 431  
 <212> DNA  
 <213> Homo sapien

<400> 360  
 aaaaaaaaaa agccagaaca acatgtgata gataatatga ttggctgcac acttccagac 60  
 tgatgaatga tgaacgtgat ggactattgt atggagcaca tcttcagcaa gagggggaaa 120  
 tactcatcat ttttggccag cagttgtttg atcaccaaac atcatgccag aatactcagc 180  
 aaaccttctt agctcttgag aagtcaaagt ccgggggaat ttattcctgg caattttaat 240  
 tggactcctt atgtgagagc agcggctacc cagctggggt ggtggagcga acccgctact 300  
 agtggacatg cagtggcaga gctcctggta accacctaga ggaatacaca ggcacatgtg 360  
 tgatgccaag cgtgacacct gtagcactca aatttgtctt gtttttgtct ttcgggtgtgt 420  
 agattcttag t 431

<210> 361  
 <211> 351  
 <212> DNA  
 <213> Homo sapien

<400> 361  
 aacttgattt ccgatcaaaa gaatcatcat ctttaccttg acttttcagg gaattactga 60  
 actttcttct cagaagatag ggcacagcca ttgccttggc ctcaacttgaa gggctctgcat 120  
 ttgggtcctc tggctctctt ccaagtttcc cagccactcg agggagaaat atcgggaggt 180  
 ttgacttcct ccggggcttt cccgagggtc tcaccgtgag ccctgcggcc ctcagggtctg 240  
 caatcctgga ttcaatgtct gaaacctcgc tctctgcctg ctggacttct gagggcgtca 300  
 ctgccactct gtctccagc tctgacagct cctcatctgt ggtcctgttg t 351

<210> 362  
 <211> 463  
 <212> DNA  
 <213> Homo sapien

<400> 362  
 acttcatcag gccataatgg gtgcctcccg tgagaatcca agcacctttg gactgcgcga 60  
 tgtagatgag ccggtgaag atcttgcgca tgcgcgggtt cagggcgaag ttcttggcgc 120  
 ccccggtcac agaaatgacc aggttgggtg ttttcagggt ccagtgtctg gtcagcagct 180  
 cgtaaaggat ttccgcgtcc gtgtcgcagg acagaogtat atacttcctt ttcttcccca 240  
 gtgtctcaaa ctgaatatcc ccaaaggcgt cggtaggaaa ttcttgggtg tgtttcttgt 300  
 agttccattt ctcacttttg ttgatctggg tgccttccat gtgctggctc tgggcatagc 360  
 cacacttgca cacattctcc ctgataagca cgatgggtgt gacaggaagg aaggatttca 420  
 ttgagcctgc ttatggaaac tggattgtt agcttaaata gac 463

<210> 363  
 <211> 653  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(653)



<400> 363

<210> 364

<211> 401

<212> DNA

<213> Homo sapien

<400> 364

actagaggaa	agacgtttaa	ccactctact	accacttggtg	gaactctcaa	agggtaaattg	60
acaaagccaa	tgaatgactc	taaaaacaat	atttacattt	aatgggtttgt	agacaataaaa	120
aaaaaaggt	ggatagatct	agaattgtaa	cattttaaga	aaaccatagc	atttgacaga	180
tgagaaaagct	caattataga	tgcaaagtta	taactaaact	actatagtag	taaagaaata	240
catttcacac	ccttcataata	aattcactat	cttggcttga	ggcactccat	aaaatgtatc	300
acgtgcatag	taaatcttta	tatttgctat	ggcgttgcac	tagaggactt	ggactgcaac	360
aagtggatgc	gcggaaaatg	aaatcttctt	caatagccca	g		401

<210> 365

<211> 356

<212> DNA

<213> Homo sapien

<400> 365

cagatgtcat	atttgggctt	aaaattttcaa	gaagggcact	tcaaatggct	ttgcatttgc	60
atgtttcagt	gctagagcgt	aggaatagac	cctggcgctc	actgtgagat	gttcttcagc	120
taccagagca	tcaagtctct	gcagcaggtc	attcttgggt	aaagaaatga	cttccacaaa	180
ctctccatcc	cctggctttg	gcttcggcct	tgcgttttcg	gcacatctc	cgттаатggт	240
gactgtcacg	atgtgtatag	tacagtttga	caagcctggg	tccatacaga	ccgctggaga	300
acattcgгca	atgtccctt	tgtagccagt	ttctttcttcg	agctcccгga	gagcag	356

<210> 366

<211> 1851

<212> DNA

<213> Homo sapien

<400> 366

```

tcattccat  tgccagcagc  ggcaccgtta  gtcaggtttt  ctgggaatcc  cacatgagta      60
cttccgtgtt  cttcattcct  cttcaatagc  cataaatcct  ctagtctctgg  ctggctgttt     120
tactttcctt  taagcctttg  tgactcttcc  tctgatgtca  gctttaagtc  ttgtttctgga    180

```







<400> 368

<210> 369

<212> DNA

<400> 369

gggtcgccca	ggggsgcgt	gggctttcct	gggtgggtg	tgggttttcc	ctgggtgggg	60
tgggctgggc	trgaatcccc	tgctggggtt	ggcaggtttt	ggctgggatt	gaacttttytc	120
ttcaaacaga	ttggaaccc	ggagttacct	gctagtgtgt	gaaactgggt	ggtagacgcg	180
atctgttggc	tactactggc	ttctcctggc	tgtaaaagc	agatgggtgt	tgaggttgat	240
tccatgcccg	ctgctttctt	tgtgaagaag	ccatttggtc	tcaggagcaa	gatgggcaag	300
tgggtgctgc	gttgcttccc	ctgctgcagg	gagagcggca	agagcaacgt	gggcacttct	360
ggagaccacg	acgactctgc	tatgaagaca	ctcaggagca	agatgggcaa	gtggtgccgc	420
cactgcttcc	cctgctgcag	ggggagtggc	aagagcaacg	tgggcgcttc	tggagaccac	480
gacgaytctg	ctatgaagac	actcaggaac	aagatgggca	agtgggtgctg	ccactgcttc	540
ccctgctgca	gggggagcrg	caagagcaag	gtgggcgctt	ggggagacta	cgatgacagy	600
gccttcatgg	akcccaggta	ccacgtccrt	ggagaagatc	tggacaagct	ccacagagct	660
gcctggtggg	gtaaagtccc	cagaaaggat	ctcatcgtca	tgctcaggga	cackgaygtg	720
aacaagargg	acaagcaaaa	gaggactgct	ctacatctgg	cctctgccaa	tgggaattca	780
gaagtagtaa	aactcstgct	ggacagacga	tgtcaactta	atgtccttga	caacaaaaag	840
aggacagctc	tgayaaaggc	cgtacaatgc	caggaagatg	aatgtgcggt	aatgttgctg	900







```
<210> 371
<211> 1855
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc_feature  
<222> (1)...(1855)  
<223> n = A,T,C or G
```

<400> 371						
tgcacgcac	ggccagtgct	tgtgccacgt	acactgagcg	ccccctgagat	gtgcacgccc	60
cacgcgcacg	ttgcacgcgc	ggcagcggt	tggctggctt	gtaacggctt	gcacgcgcac	120
gcgcgcgcgc	cataaccgtc	agactggcct	gtaacggctt	gcaggcgcac	gccgcacgcg	180
cgtaacggct	tggctgcctt	gtaacggctt	gcacgtgcat	gctgcacgcg	cgttaacggc	240
ttggctggca	tgtagccgct	tggcttggct	ttgcattytt	tgetkggctk	ggcgttgkty	300
tcttgattg	acgcttctct	cttggaatkga	cgtttctctc	ttggatkgac	gtttctytyty	360
tgcgcttctt	ttgctggact	tgacctttty	tctgctgggt	ttggcattcc	tttgggggtgg	420
gctgggtggt	ttctccgggg	gggktkgccc	ttctgggggt	gggcgtgggk	cgcccccagg	480
gggcgtgggc	tttccccggg	tggtgtggg	ttttctggg	gtgggggtggg	ctgtgctggg	540
atccccctgc	tggggttggc	agggattgac	ttttttcttc	aaacagattg	gaaacccgga	600
gtaacntgct	agttggtgaa	actggttggg	agacgcgac	tgetggtact	actgtttctc	660
ctggctgtta	aaagcagatg	gtggctgagg	ttgattcaat	gcgggtgct	tcttctgtga	720
agaagccatt	tggctctcagg	agcaagatgg	gcaagtgggtg	cgccactgct	tccccctgctg	780
cagggggagc	ggcaagagca	acgtgggcac	ttctggagac	cacaacgact	cctctgtgaa	840
gacgcttggg	agcaagaggt	gcaagtgggtg	ctgccactg	cttccccctgc	tgcagggggag	900
cggcaagagc	aacgtggkcg	cttggggaga	ctacgatgac	agcgcttcca	tggakcccag	960
gtaccacgtc	crtggagaag	atctggacaa	gctccacaga	gctgctgggt	ggggtaaagt	1020
ccccagaaag	gatctcatcg	tcatgctcag	ggacactgay	gtgaacaaga	rggacaagca	1080
aaagaggact	gctctacatc	tggcctctgc	caatgggaat	tcagaagtag	taaaactcgt	1140
gctggacaga	cgatgtcaac	ttaatgtcct	tgacaacaaa	aagaggacag	ctctgacaaa	1200
ggcgtacaa	tgccaggaag	atgaatgtgc	gttaatgttg	ctggaacatg	gcactgatcc	1260
aaatattcca	gatgagtatg	gaaataccac	tctacactat	gctgtctaca	atgaagataa	1320
attaatggcc	aaagcactgc	tcttatacgg	tgctgatatc	gaatcaaaaa	acaagggtata	1380
gatctactaa	ttttatcttc	aaaatactga	aatgcattca	ttttaacatt	gacgtgtgta	1440
agggccagtc	ttcgtatttt	ggaagctcaa	gcataaacttg	aatgaaaata	ttttgaaatg	1500
acctaattat	ctaagacttt	attttaaata	ttgttatttt	caaagaagca	ttagagggta	1560
cagttttttt	tttttaaattg	cacttctgggt	aaatactttt	gttgaaaaca	ctgaatttgt	1620
aaaaggtaat	acttactatt	tttcaatttt	tcctctctag	gattttttttc	ccctaattgaa	1680
tgtaagatgg	caaaatttgc	cctgaaatag	gttttacatg	aaaactccaa	gaaaagttaa	1740
acatgtttca	gtgaatagag	atcctgctcc	tttggaagt	tcctaaaaaa	cagtaataga	1800



<400> 372

<400> 373

atggtggttg	aggttgattc	catgccggct	gcctcttctg	tgaagaagcc	atttgggtctc	60
aggagcaaga	tgggcaagtg	gtgctgccgt	tgttccccct	gctgcaggga	gagcggcaag	120
agcaacgtgg	gcacttctgg	agaccacgac	gactctgcta	tgaagacact	caggagcaag	180
atgggcaagt	ggtgccgcca	ctgcttcccc	tgtctcaggg	ggagtggcaa	gagcaacgtg	240
ggcgcttctg	gagaccacga	cgactctgct	atgaagacac	tcaggaacaa	gatgggcaag	300
tggtgctgcc	actgcttccc	ctgctgcagg	gggagcggca	agagcaagggt	gggcgcttgg	360
ggagactacg	atgacagtgc	cttcattggag	cccaggtacc	acgtccgtgg	agaagatctg	420
gacaagctcc	acagagctgc	ctggtgggggt	aaagtcccca	gaaaggatct	catcgtcatg	480
ctcaggggaca	ctgacgtgaa	caagaaggac	aagcaaaaga	ggactgctct	acatctggcc	540
tctgccaatg	ggaattcaga	agtagtaaaa	ctctgctggg	acagacgatg	tcaacttaat	600
gtccttgaca	acaaaaagag	gacagctctg	ataaaggccg	tacaatgcca	ggaagatgaa	660
tgtgcgttaa	tgttgctgga	acatggcact	gatccaaata	ttccagatga	gtatggaaat	720
accactctgc	actacgtat	ctataatgaa	gataaattaa	tggccaaagc	actgctctta	780
tatggtgctg	atatcgaatc	aaaaaacaag	catggcctca	caccactggt	acttggtgta	840
catgagcaaaa	aacagcaagt	cgtgaaattt	ttaatcaaga	aaaaagcgaa	tttaaattgca	900
ctggaatagat	atggaaggac	tgtctctcata	cttgcgtgat	gttggtggatc	agcaagtata	960
gtcagccttc	tacttgagca	aaatattgat	gtatcttctc	aagatctatc	tggacagacg	1020
gccagagagt	atgctgtttc	tagtcatcat	catgtaattt	gccagttaact	ttctgactac	1080



<400> 374

<400> 375

atggtggttg aggttgattc catgccggct gcctcttctg tgaagaagcc atttggtctc 60



```

aggagcaaga tgggcaagtg gtgctgccgt tgcttcccct gctgcagggg gagcggcaag 120
agcaacgtgg gcacttctgg agaccacgac gactctgcta tgaagacact caggagcaag 180
atgggcaagt ggtgccgcca ctgcttcccc tgctgcaggg ggagtggcaa gagcaacgtg 240
ggcgcttctg gagaccacga cgactctgct atgaagacac tcaggaacaa gatgggcaag 300
tgggtgctgcc actgcttccc ctgctgcagg gggagcggca agagcaaggt gggcgcttgg 360
ggagactacg atgacagtgc cttcatggag ccaggtacc acgtccgtgg agaagatctg 420
gacaagctcc acagagctgc ctggtggggg aaagtcccca gaaaggatct catcgtcatg 480
ctcagggaca ctgacgtgaa caagaaggac aagcaaaaga ggactgctct acatctggcc 540
tctgccaatg ggaattcaga agtagtaaaa ctctgctgg acagacgatg tcaacttaat 600
gtccttgaca aaaaaaagag gacagctctg ataaaggccg tacaatgcca ggaagatgaa 660
tgtgcgttaa tgttgctgga acatggcact gatccaaata ttccagatga gtatggaaat 720
accactctgc actacgctat ctataatgaa gataaattaa tggccaaagc actgctctta 780
tatggtgctg atatcgaatc aaaaaacaag catggcctca caccactgtt acttgggtgta 840
catgagcaaa aacagcaagt cgtgaaattt ttaatcaaga aaaaagcgaa tttaaatgca 900
ctggatagat atggaaggac tgctctcata cttgctgtat gttgtggatc agcaagtata 960
gtcagccttc tacttgagca aaatattgat gtatcttctc aagatctatc tggacagacg 1020
gccagagagt atgctgtttc tagtcatcat catgtaatth gccagttact ttctgactac 1080
aaagaaaaac agatgctaaa aatctcttct gaaaacagca atccagaaca agacttaaag 1140
ctgacatcag aggaagagtc acaaaggttc aaaggcagtg aaaatagcca gccagagaaa 1200
atgtctcaag aaccagaaat aaataaggat ggtgatagag aggttgaaga agaaatgaag 1260
aagcatgaaa gtaataatgt gggattacta gaaaacctga ctaatggtgt cactgctggc 1320
aatggtgata atggattaat tctcaaaagg aagagcagaa cacctgaaaa tcagcaatth 1380
cctgacaacg aaagtgaaga gtatcacaga atttgcgaat tagtttctga ctacaaagaa 1440
aaacagatgc caaaatactc ttctgaaaac agcaaccacg aacaagactt aaagctgaca 1500
tcagaggaag agtcacaaag gcttgagggc agtgaaaatg gccagccaga gaaaagatct 1560
caagaaccag aaataaataa ggatggtgat agagagctag aaaatthtat ggctatcgaa 1620
gaaatgaaga agcacggaag tactcatgtc ggattcccag aaaacctgac taatggtgcc 1680
actgctggca atggtgatga tggattaatt cctccaagga agagcagaac acctgaaagc 1740
cagcaatthc ctgacactga gaatgaagag tatcacagtg acgaacaaaa tgatactcag 1800
aagcaattht gtgaagaaca gaacactgga atattacacg atgagattct gattcatgaa 1860
gaaaagcaga tagaagtggg tgaaaaaatg aattctgagc tttctcttag ttgtaagaaa 1920
gaaaagacaa tcttgcagta aaatagtacg ttgcgggaag aaattgccat gctaagactg 1980
gagctagaca caatgaaaca tcagagccag ctaaaaaaa aaaaaaaaaa aaaaaaaaaa 2040

```

<210> 376

<211> 329

<212> PRT

<213> Homo sapien

<400> 376

```

Met Asp Ile Val Val Ser Gly Ser His Pro Leu Trp Val Asp Ser Phe
 1             5             10            15
Leu His Leu Ala Gly Ser Asp Leu Leu Ser Arg Ser Leu Met Ala Glu
                20            25            30
Glu Tyr Thr Ile Val His Ala Ser Phe Ile Ser Cys Ile Ser Ser Ser
                35            40            45
Leu Asp Gly Gln Gly Glu Arg Gln Glu Gln Arg Gly His Phe Trp Arg
                50            55            60
Pro Gln Arg Leu Leu Cys Glu Asp Ala Trp Glu Gln Glu Val Gln Val
        65              70              75              80
Val Leu Pro Leu Leu Pro Leu Leu Gln Gly Ser Gly Lys Ser Asn Val

```



				85					90				95			
Val	Ala	Trp	Gly	Asp	Tyr	Asp	Asp	Ser	Ala	Phe	Met	Asp	Pro	Arg	Tyr	
			100					105					110			
His	Val	His	Gly	Glu	Asp	Leu	Asp	Lys	Leu	His	Arg	Ala	Ala	Trp	Trp	
		115					120					125				
Gly	Lys	Val	Pro	Arg	Lys	Asp	Leu	Ile	Val	Met	Leu	Arg	Asp	Thr	Asp	
	130					135					140					
Val	Asn	Lys	Arg	Asp	Lys	Gln	Lys	Arg	Thr	Ala	Leu	His	Leu	Ala	Ser	
145					150					155					160	
Ala	Asn	Gly	Asn	Ser	Glu	Val	Val	Lys	Leu	Val	Leu	Asp	Arg	Arg	Cys	
			165					170						175		
Gln	Leu	Asn	Val	Leu	Asp	Asn	Lys	Lys	Arg	Thr	Ala	Leu	Thr	Lys	Ala	
		180					185						190			
Val	Gln	Cys	Gln	Glu	Asp	Glu	Cys	Ala	Leu	Met	Leu	Leu	Glu	His	Gly	
	195						200					205				
Thr	Asp	Pro	Asn	Ile	Pro	Asp	Glu	Tyr	Gly	Asn	Thr	Thr	Leu	His	Tyr	
	210					215					220					
Ala	Val	Tyr	Asn	Glu	Asp	Lys	Leu	Met	Ala	Lys	Ala	Leu	Leu	Leu	Tyr	
225				230						235					240	
Gly	Ala	Asp	Ile	Glu	Ser	Lys	Asn	Lys	His	Gly	Leu	Thr	Pro	Leu	Leu	
			245						250					255		
Leu	Gly	Ile	His	Glu	Gln	Lys	Gln	Gln	Val	Val	Lys	Phe	Leu	Ile	Lys	
		260					265						270			
Lys	Lys	Ala	Asn	Leu	Asn	Ala	Leu	Asp	Arg	Tyr	Gly	Arg	Thr	Ala	Leu	
	275						280					285				
Ile	Leu	Ala	Val	Cys	Cys	Gly	Ser	Ala	Ser	Ile	Val	Ser	Pro	Leu	Leu	
	290					295					300					
Glu	Gln	Asn	Val	Asp	Val	Ser	Ser	Gln	Asp	Leu	Glu	Arg	Arg	Pro	Glu	
305				310						315					320	
Ser	Met	Leu	Phe	Leu	Val	Ile	Ile	Met								
				325												

&lt;210&gt; 377

&lt;211&gt; 148

&lt;212&gt; PRT

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; VARIANT

&lt;222&gt; (1)...(148)

&lt;223&gt; Xaa = Any Amino Acid

&lt;400&gt; 377

Met	Thr	Xaa	Pro	Ser	Trp	Ser	Pro	Gly	Thr	Thr	Ser	Val	Glu	Lys	Ile	
1				5				10					15			
Trp	Thr	Ser	Ser	Thr	Glu	Leu	Pro	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys	
		20					25					30				
Asp	Leu	Ile	Val	Met	Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Xaa	Asp	Lys	
	35					40					45					
Gln	Lys	Arg	Thr	Ala	Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu	
50						55					60					







Ala	Leu	Leu	Leu	Tyr	Gly	Ala	Asp	Ile	Glu	Ser	Lys	Asn	Lys	His	Gly
			260					265					270		
Leu	Thr	Pro	Leu	Leu	Leu	Gly	Val	His	Glu	Gln	Lys	Gln	Gln	Val	Val
			275					280					285		
Lys	Phe	Leu	Ile	Lys	Lys	Lys	Ala	Asn	Leu	Asn	Ala	Leu	Asp	Arg	Tyr
			290					295				300			
Gly	Arg	Thr	Ala	Leu	Ile	Leu	Ala	Val	Cys	Cys	Gly	Ser	Ala	Ser	Ile
305					310					315					320
Val	Ser	Leu	Leu	Leu	Glu	Gln	Asn	Ile	Asp	Val	Ser	Ser	Gln	Asp	Leu
				325					330					335	
Ser	Gly	Gln	Thr	Ala	Arg	Glu	Tyr	Ala	Val	Ser	Ser	His	His	His	Val
			340					345					350		
Ile	Cys	Gln	Leu	Leu	Ser	Asp	Tyr	Lys	Glu	Lys	Gln	Met	Leu	Lys	Ile
			355					360					365		
Ser	Ser	Glu	Asn	Ser	Asn	Pro	Glu	Asn	Val	Ser	Arg	Thr	Arg	Asn	Lys
			370				375					380			
Pro	Arg	Thr	His	Met	Val	Val	Glu	Val	Asp	Ser	Met	Pro	Ala	Ala	Ser
385					390					395					400
Ser	Val	Lys	Lys	Pro	Phe	Gly	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp	Cys
				405					410					415	
Cys	Arg	Cys	Phe	Pro	Cys	Cys	Arg	Glu	Ser	Gly	Lys	Ser	Asn	Val	Gly
			420					425					430		
Thr	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Ser	Lys
			435					440					445		
Met	Gly	Lys	Trp	Cys	Arg	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser	Gly
			450				455					460			
Lys	Ser	Asn	Val	Gly	Ala	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys
465					470					475					480
Thr	Leu	Arg	Asn	Lys	Met	Gly	Lys	Trp	Cys	Cys	His	Cys	Phe	Pro	Cys
				485					490					495	
Cys	Arg	Gly	Ser	Gly	Lys	Ser	Lys	Val	Gly	Ala	Trp	Gly	Asp	Tyr	Asp
			500					505	-				510		
Asp	Ser	Ala	Phe	Met	Glu	Pro	Arg	Tyr	His	Val	Arg	Gly	Glu	Asp	Leu
			515					520					525		
Asp	Lys	Leu	His	Arg	Ala	Ala	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys	Asp
			530				535					540			
Leu	Ile	Val	Met	Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Lys	Asp	Lys	Gln
545					550					555					560
Lys	Arg	Thr	Ala	Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu	Val
				565					570					575	
Val	Lys	Leu	Leu	Leu	Asp	Arg	Arg	Cys	Gln	Leu	Asn	Val	Leu	Asp	Asn
			580					585					590		
Lys	Lys	Arg	Thr	Ala	Leu	Ile	Lys	Ala	Val	Gln	Cys	Gln	Glu	Asp	Glu
			595				600					605			
Cys	Ala														







Gly	Lys	Trp	Cys	Cys	Arg	Cys	Phe	Pro	Cys	Cys	Arg	Glu	Ser	Gly	Lys
1075				1080				1085							
Ser	Asn	Val	Gly	Thr	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys	Thr
1090				1095				1100							
Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp	Cys	Arg	His	Cys	Phe	Pro	Cys	Cys
1105				1110				1115				112			
Arg	Gly	Ser	Gly	Lys	Ser	Asn	Val	Gly	Ala	Ser	Gly	Asp	His	Asp	Asp
1125				1130				1135							
Ser	Ala	Met	Lys	Thr	Leu	Arg	Asn	Lys	Met	Gly	Lys	Trp	Cys	Cys	His
1140				1145				1150							
Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser	Gly	Lys	Ser	Lys	Val	Gly	Ala	Trp
1155				1160				1165							
Gly	Asp	Tyr	Asp	Asp	Ser	Ala	Phe	Met	Glu	Pro	Arg	Tyr	His	Val	Arg
1170				1175				1180							
Gly	Glu	Asp	Leu	Asp	Lys	Leu	His	Arg	Ala	Ala	Trp	Trp	Gly	Lys	Val
1185				1190				1195				120			
Pro	Arg	Lys	Asp	Leu	Ile	Val	Met	Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys
1205				1210				1215							
Lys	Asp	Lys	Gln	Lys	Arg	Thr	Ala	Leu	His	Leu	Ala	Ser	Ala	Asn	Gly
1220				1225				1230							
Asn	Ser	Glu	Val	Val	Lys	Leu	Leu	Leu	Asp	Arg	Arg	Cys	Gln	Leu	Asn
1235				1240				1245							
Val	Leu	Asp	Asn	Lys	Lys	Arg	Thr	Ala	Leu	Ile	Lys	Ala	Val	Gln	Cys
1250				1255				1260							
Gln	Glu	Asp	Glu	Cys	Ala	Leu	Met	Leu	Leu	Glu	His	Gly	Thr	Asp	Pro
1265				1270				1275				128			
Asn	Ile	Pro	Asp	Glu	Tyr	Gly	Asn	Thr	Thr	Leu	His	Tyr	Ala	Ile	Tyr
1285				1290				1295							
Asn	Glu	Asp	Lys	Leu	Met	Ala	Lys	Ala	Leu	Leu	Leu	Tyr	Gly	Ala	Asp
1300				1305				1310							
Ile	Glu	Ser	Lys	Asn	Lys	His	Gly	Leu	Thr	Pro	Leu	Leu	Leu	Gly	Val
1315				1320				1325							
His	Glu	Gln	Lys	Gln	Gln	Val	Val	Lys	Phe	Leu	Ile	Lys	Lys	Lys	Ala
1330				1335				1340							
Asn	Leu	Asn	Ala	Leu	Asp	Arg	Tyr	Gly	Arg	Thr	Ala	Leu	Ile	Leu	Ala
1345				1350				1355				136			
Val	Cys	Cys	Gly	Ser	Ala	Ser	Ile	Val	Ser	Leu	Leu	Leu	Glu	Gln	Asn
1365				1370				1375							
Ile	Asp	Val	Ser	Ser	Gln	Asp	Leu	Ser	Gly	Gln	Thr	Ala	Arg	Glu	Tyr
1380				1385				1390							
Ala	Val	Ser	Ser	His	His	His	Val	Ile	Cys	Gln	Leu	Leu	Ser	Asp	Tyr
1395				1400				1405							
Lys	Glu	Lys	Gln	Met	Leu	Lys	Ile	Ser	Ser	Glu	Asn	Ser	Asn	Pro	Glu
1410				1415				1420							
Gln	Asp	Leu	Lys	Leu	Thr	Ser	Glu	Glu	Glu	Ser	Gln	Arg	Phe	Lys	Gly
1425				1430				1435				144			
Ser	Glu	Asn	Ser	Gln	Pro	Glu	Lys	Met	Ser	Gln	Glu	Pro	Glu	Ile	Asn
1445				1450				1455							
Lys	Asp	Gly	Asp	Arg	Glu	Val	Glu	Glu	Glu	Met	Lys	Lys	His	Glu	Ser
1460				1465				1470							
Asn	Asn	Val	Gly	Leu	Leu	Glu	Asn	Leu	Thr	Asn	Gly	Val	Thr	Ala	Glu











																<210>	380
																<211>	671
																<212>	PRT
																<213>	Homo sapien
																<400>	380
Met	Val	Val	Glu	Val	Asp	Ser	Met	Pro	Ala	Ala	Ser	Ser	Val	Lys	Lys		
1				5					10					15			
Pro	Phe	Gly	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp	Cys	Cys	Arg	Cys	Phe		
			20					25					30				
Pro	Cys	Cys	Arg	Glu	Ser	Gly	Lys	Ser	Asn	Val	Gly	Thr	Ser	Gly	Asp		
		35					40					45					
His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp		
	50					55					60						
Cys	Arg	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser	Gly	Lys	Ser	Asn	Val		
65					70				75					80			
Gly	Ala	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Asn		
				85					90					95			
Lys	Met	Gly	Lys	Trp	Cys	Cys	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser		
			100					105					110				
Gly	Lys	Ser	Lys	Val	Gly	Ala	Trp	Gly	Asp	Tyr	Asp	Asp	Ser	Ala	Phe		
		115					120					125					
Met	Glu	Pro	Arg	Tyr	His	Val	Arg	Gly	Glu	Asp	Leu	Asp	Lys	Leu	His		
	130					135					140						
Arg	Ala	Ala	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys	Asp	Leu	Ile	Val	Met		
145					150					155					160		
Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Lys	Asp	Lys	Gln	Lys	Arg	Thr	Ala		
				165					170					175			
Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu	Val	Val	Lys	Leu	Leu		
			180					185					190				
Leu	Asp	Arg	Arg	Cys	Gln	Leu	Asn	Val	Leu	Asp	Asn	Lys	Lys	Arg	Thr		
		195					200					205					







610		615		620	
Glu Val Val Glu Lys Met Asn Ser Glu Leu Ser Leu Ser Cys Lys Lys					
625		630		635	640
Glu Lys Asp Ile Leu His Glu Asn Ser Thr Leu Arg Glu Glu Ile Ala					
	645		650		655
Met Leu Arg Leu Glu Leu Asp Thr Met Lys His Gln Ser Gln Leu					
660		665		670	

<210> 381  
 <211> 251  
 <212> DNA  
 <213> Homo sapien

<400> 381	
ggagaagcgt ctgctggggc aggaaggggt ttccctgccc tctcacctgt ccctcaccaa	60
ggtaacatgc ttcccctaag ggtatcccaa cccagggggc tcaccatgac ctctgagggg	120
ccaatatccc aggagaagca ttggggaggt gggggcaggt gaaggacca ggactcacac	180
atcctggggc tccaaggcag aggagaggggt cctcaagaag gtcaggagga aaatccgtaa	240
caagcagtca g	251

<210> 382  
 <211> 3279  
 <212> DNA  
 <213> Homo sapiens

<400> 382	
cttcctgcag ccccatgct ggtgaggggc acgggagga acagtggacc caacatggaa	60
atgctggagg gtgtcaggaa gtgatcgggc tctggggcag ggaggagggg tggggagtgt	120
cactgggagg ggacatcctg cagaaggtag gagtgcagca acaccgctg caggggaggg	180
gagagccctg cggcacctg gggagcagag ggagcagcac ctgcccaggc ctgggaggag	240
gggcctggag ggcgtgagga ggagcgaggg ggctgcatgg ctggagttag ggatcagggg	300
cagggcgcca gatggcctca cacagggaag agagggcccc tctgcaggg cctcacctgg	360
gccacaggag gacactgctt ttctctgag gagtgcaggag ctgtggatgg tgctggacag	420
aagaaggaca gggcctggct cagggtgtcca gaggtgtcg ctggcttccc tttgggatca	480
gactgcaggg agggagggcg gcagggttgt ggggggagtg acgatgagga tgacctgggg	540
gtggctccag gccttgcctc tgctggggc ctcaccagc ctccctcaca gtctcctggc	600
cctcagtctc tccctccac tccatcctcc atctggcctc agtgggtcat tctgatcact	660
gaactgacca taccagccc tgcccaaggc cctccatggc tccccaatgc cctggagagg	720
ggacatctag tcagagagta gtcctgaaga ggtggcctct gcgatgtgcc tgtgggggca	780
gcacatctga gatgtcccg gccctcatcc tgctgacctg tctgcaggga ctgtcctcct	840
ggaccttgcc ccttgtgcag gagctggacc ctgaagtccc ctccccatag gccaagactg	900
gagccttggt cctctgttg gactccctgc ccatattctt gtgggagtgg gttctggaga	960
catttctgtc tgttctgag agctgggaat tgctctcagt catctgcctg cgcgggtctg	1020
agagatggag ttgcctaggc agttattggg gccaatctt ctcacttgtt ctctcctcct	1080
ttacccttag ggtgattctg ggggtccact tgtctgtaat ggtgtgcttc aaggtatcac	1140
atcatggggc cctgagccat gtgccctgcc tgaaaagcct gctgtgtaca ccaaggtggg	1200
gcattaccgg aagtggatca aggacaccat cgcagccaac ccctgagtgc ccctgtccca	1260
cccctacctc tagtaaattt aagtccacct caggttctgg catcacttgg ctttcttgga	1320
tgctggacac ctgaagcttg gaactcacct ggccgaagct ctagcctcct gagtctact	1380
gacctgtgct ttctgggtgt gagtccaggg ctgctaggaa aaggaatggg cagacacagg	1440
tgtatgcca tgtttctgaa atgggtataa tttcgtcctc tccttcggaa cactggctgt	1500



```
<210> 383
<211> 155
<212> PRT
<213> Homo sapiens

<400> 383
Met Ala Gly Val Arg Asp Gln Gly Gln Gly Ala Arg Trp Pro His Thr
          5                                10                        15

Gly Lys Arg Gly Pro Leu Leu Gln Gly Leu Thr Trp Ala Thr Gly Gly
          20                                25                        30

His Cys Phe Ser Ser Glu Glu Ser Gly Ala Val Asp Gly Ala Gly Gln
          35                                40                        45

Lys Lys Asp Arg Ala Trp Leu Arg Cys Pro Glu Ala Val Ala Gly Phe
          50                                55                        60

Pro Leu Gly Ser Asp Cys Arg Glu Gly Gly Arg Gln Gly Cys Gly Gly
          65                                70                        75                        80
```







cggtgcccc	gtttgacaga	aggaaaggcg	gagcttattc	aaagtctaga	gggagtgagg	60
gagttaaggc	tggatttcag	atctgcctgg	ttccagccgc	agtgtgccct	ctgctcccc	120
aacgactttc	caaataatct	caccagcgcc	ttccagctca	ggcgctctag	aagcgtcttg	180
aagcctatgg	ccagctgtct	ttgtgttccc	tctcacccgc	ctgtcctcac	agctgagact	240
cccaggaaac	cttcagacta	ccttcctctg	ccttcagcaa	ggggcggttg	ccacattctc	300



tgagggtcag tggaagaacc tagactccca ttgctagagg tagaaagggg aaggggtgctg 360  
gggag 365

<210> 390  
<211> 221  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> (1)...(221)  
<223> n = A,T,C or G

<400> 390  
tgccctctcca tcctggcccc gacttctctg tcaggaaagt ggggatggac cccatctgca 60  
tacacggntt ctcctgggtg tggaacatct ctgcttgagg tttcaggaag gcctctggct 120  
gctctangag tctgancnga ntcgttgccc cantntgaca naaggaaagg cggagcttat 180  
tcaaagtcta gaggagtgagg aggagttaag gctggatttc a 221

<210> 391  
<211> 325  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> (1)...(325)  
<223> n = A,T,C or G

<400> 391  
tggagcaggt cccgaggcct ccctagagcc tggggccgac tctgtgncga tgcangcttt 60  
ctctcgcgcc cagcctggag ctgctcctgg catctacca caatcagnog aggcgagcag 120  
tagccagggc actgctgcca acagccagtc cnnataccat catgtnaccc ggtgngctct 180  
naanttngat ntccanagcc ctaccccatn tagttctgct cccccaccgg ntaccagccc 240  
cactgcccag gaatcctaca gccagtaccc tgtcccgacg tctctaccta ccagtacgat 300  
gagacctccg gctactacta tgacc 325

<210> 392  
<211> 277  
<212> DNA  
<213> Homo sapiens

<220>  
<221> misc\_feature  
<222> (1)...(277)  
<223> n = A,T,C or G

<400> 392  
atattgttta actccttcct ttatatcttt taacattttc atggngaaaag gttcacatct 60  
agtctcactt nggcnagnn ctcctacttg agtctcttcc cgggctggn ccagtngnaa 120  
antaccanga accgncatgn cttaanaacn ncctgggttn tgggttnntc aatgactgca 180



```

tgcagtgcac caccctgtcc actacgtgat gctgtaggat taaagtctca cagtgggcgg 240
ctgaggatac agcgccgcgt cctgtgttgc tggggaa 277

```

<210> 393

<211> 566

<212> DNA

<213> Homo sapiens

<400> 393

```

actagtccag tgtggtggaa ttgcgcggcg cgtcgaaggga caggtcagct gtctggctca 60
gtgatctaca ttctgaagtt gtctgaaaat gtcttcatga tttaaattcag cctaaacggt 120
ttgccgggaa cactgcagag acaatgctgt gagtttccaa ccttagccca tctgcgggca 180
gagaaggctc agtttgtcca tcagcattat catgatatac ggactgggta cttgggtaag 240
gaggggtcta ggagatctgt cccttttaga gacaccttac ttataatgaa gtatttggga 300
gggtggtttt caaaagtaga aatgtcctgt attccgatga tcatcctgta aacattttat 360
catttattaa tcatccctgc ctgtgtctat tattatattc atatctctac gctggaaact 420
ttctgcctca atgtttactg tgcctttgtt tttgctagtt tgtgttggtg aaaaaaaaaa 480
cattctctgc ctgagtttta atttttgtcc aaagttattt taatctatac aattaaaagc 540
ttttgcctat caaaaaaaaa aaaaaa 566

```

<210> 394

<211> 384

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(384)

<223> n = A,T,C or G

<400> 394

```

gaacatacat gtcccggcac ctgagctgca gtctgacatc atcgccatca cgggcctcgc 60
tgcaaattng gaccgggcca aggctggact gctggagcgt gtgaaggagc tacaggccna 120
gcaggaggac cgggctttaa ggagttttaa gctgagtgtc actgtagacc ccaaatacca 180
tccaagatt atcgggagaa agggggcagt aattacccaa atcgggttg agcatgacgt 240
gaacatccag tttcctgata aggacgatgg gaaccagccc caggaccaa ttaccatcac 300
agggtagcaa aagaacacag aagctgccag ggatgctata ctgagaattg tgggtgaact 360
tgagcagatg gtttctgagg acgt 384

```

<210> 395

<211> 399

<212> DNA

<213> Homo sapiens

<400> 395

```

ggcaaaactg tgtgacctca ataagacctc gcagatccaa ggtcaagtat cagaagtgc 60
tctgaccttg gactccaaga cctacatcaa cagcctggct atattagatg atgagccagt 120
tatcagaggt ttcatcattg cggaaattgt ggagtctaag gaaatcatgg cctctgaagt 180
attcacgtct ttccagtacc ctgagttctc tatagagttg cctaacacag gcagaattgg 240
ccagctactt gtctgcaatt gtatcttcaa gaataccctg gccatccctt tgactgacgt 300
caagttctct ttggaaagcc tgggcatctc ctactacag acctctgacc atgggacggt 360

```



399

<400>	396					
tggagttntc	agtgcaaaca	agccataaag	cttcagtagc	aaattactgt	ctcacagaaa	60
gacattttca	actttctgctc	cagctgctga	taaaacaaat	catgtgttta	gottgactcc	120
agacaaggac	aacctgttc	ttcataactc	tctagagaaa	aaaaggagtt	gttagtagat	180
actaaaaaaaa	gtggatgaat	aatctggata	tttttcttaa	aaagattcct	tgaaacacat	240
taggaaaaatg	gagggcctta	tgatcagaat	gctagaatta	gtccattgtg	ctgaagcagg	300
gtttagggga	gggagtgagg	gataaaagaa	ggaaaaaaag	aagagtgaga	aaacctattt	360
atcaaagcag	gtgctatcac	tcaatgttag	gcctgtctct	ttt		403

```
<400> 397
actagtncag tgtggtggaa ttgcgggccg cgtcgaccta naanccatct ctatagcaaa 60
tccatccccg ctctctggttg gtnacagaat gactgacaaa 100
```

```
<220>
<221> misc_feature
<222> (1)...(278)
<223> n = A,T,C or G
```

```
<400> 398
gcggccgcgt cgacagcagt tccgccagcg ctgcgccctg ggtgggggatg tgctgcacgc 60
ccacctggac atctggaagt cagcggcctg gatgaaagag cggacttcac ctggggcgat 120
tcactactgt gcctcgacca gtgaggagag ctggaccgac agcgaggtgg actcatcatg 180
ctccgggcag cccatccacc tgtggcagtt cctcaaggag ttgctactca agccccacag 240
ctatggccgc ttcattanqt qqctcaacaa qgagaagg          278
```



<210> 399  
 <211> 298  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1) ... (298)  
 <223> n = A,T,C or G

<400> 399  
 acggaggtgg aggaagcgc cctgggatcg anaggatggg tcctgncatt gaccnccctn 60  
 ggggtgccng catggagcgc atgggcgcgg gcctgggcca cggcatggat cgcgtgggct 120  
 ccgagatcga gcgcattggc ctggatcatgg accgcatggg ctccgtggag cgcattgggct 180  
 ccggcattga gcgcattggc ccgctggggc tcgaccacat ggccctccanc attgancgca 240  
 tgggccagac catggagcgc attggctctg gcgtggagcn catgggtgcc ggcattggg 298

<210> 400  
 <211> 548  
 <212> DNA  
 <213> Homo sapiens

<400> 400  
 acatcaacta cttcctcatt ttaaggatcg gcagttccct tcctccctctt ttctgcctt 60  
 gtacatgtac atgtatgaaa tttccttctc ttaccgaact ctctccacac atcacaagg 120  
 caaagaacca cacgcttaga agggtaagag ggcaccctat gaaatgaaat ggtgatttct 180  
 tgagtctctt ttttccacgt ttaagggggc atggcaggac ttagagttgc gagttaagac 240  
 tgcagagggc tagagaatta tttcatacag gctttgaggc caccatgtc acttatcccg 300  
 tataccctct caccatcccc ttgtctactc tgatgcccc aagatgcaac tgggcagcta 360  
 gttggcccca taattctggg cctttgttgt ttgttttaac tacttgggca tcccaggaag 420  
 ctttccagtg atctctacc atgggcccc ctctgaggat caagccctc ccaggccctg 480  
 tccccagccc ctctgcccc agcccacccg cttgccttgg tgctcagccc tcccattggg 540  
 agcaggtt 548

<210> 401  
 <211> 355  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1) ... (355)  
 <223> n = A,T,C or G

<400> 401  
 actgtttcca tggtatgttt ctacacattg ctacctcagt gctcctggaa acttagcttt 60  
 tgatgtctcc aagtagtcca ctttcattta actctttgaa actgtatcat ctttgccaag 120  
 taagagtggg ggctatttc agctgctttg acaaatgac tggctcctga cttaacgttc 180  
 tataaatgaa tgtgctgaag caaagtggc atgggtggcg cgaagaagan aaagatgtgt 240  
 tttgttttgg actctctgtg gtcccttcca atgctgnggg tttccaacca ggggaagggt 300



<210> 405



<211> 334  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(334)  
 <223> n = A,T,C or G

<400> 405  
 gagctgttat actgtgagtt ctactaggaa atcatcaaat ctgaggggtg tctggaggac 60  
 ttcaatacac ctcccccat agtgaatcag cttccagggg gtccagtccc tctccttact 120  
 tcatcccat cccatgccaa aggaagaccc tccctccttg gctcacagcc ttctctaggc 180  
 ttcccagtcg ctccaggaca gagtgggtta tgttttcagc tccatccttg ctgtgagtgt 240  
 ctgggtgcggt tgtgcctcca gcttctgctc agtgcctcat ggacagtgtc cagcccatgt 300  
 cactctccac tctctcanng tggatcccac ccct 334

<210> 406  
 <211> 216  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(216)  
 <223> n = A,T,C or G

<400> 406  
 tttcatacct aatgagggag ttganatnac atnnaaccag gaaatgcatg gatctcaang 60  
 gaaacaaaca cccaataaac tcggagtggc agactgacaa ctgtgagaca tgcacttgct 120  
 acnaaacaca aatttnatgt tgcacccttg tttctacacc tgtgggttat gacaaagaca 180  
 actgccaag aatnttcaag aaggaggact gccant 216

<210> 407  
 <211> 413  
 <212> DNA  
 <213> Homo sapiens

<400> 407  
 gctgacttgc tagtatcatc tgcattcatt gaagcacaag aacttcatgc cttgactcat 60  
 gtaaattgcaa taggattaaa aaataaattt gatatcacat ggaaacagac aaaaaatatt 120  
 gtacaacatt gcacccagtg tcagattcta cacctggcca ctgaggaagc aagagttaat 180  
 cccagaggtc tatgtcctaa tgtgttatgg caaatggatg tcatgcacgt accttcattt 240  
 ggaaaattgt catttgtcca tgtgacagtt gatacttatt cacatttcat atgggcaacc 300  
 tgccagacag gagaaagtct tcccatgtta aaagacattt attatcttgt tttcctgtca 360  
 tgggagttcc agaaaaagt taaaacagaca atgggccagg ttctgtagta aag 413

<210> 408  
 <211> 183  
 <212> DNA  
 <213> Homo sapiens



```
<210> 411
<211> 261
<212> DNA
<213> Homo sapiens
```



<220>  
 <221> misc\_feature  
 <222> (1)...(261)  
 <223> n = A,T,C or G

<400> 411  
 agagatattn cttaggtnaa agttcataga gttcccatga actatatgac tggccacaca 60  
 ggatcttttg tatttaagga ttctgagatt ttgcttgagc aggattagat aaggctgttc 120  
 tttaaatgtc tgaaatggaa cagatttcaa aaaaaaaccc cacaatctag ggtgggaaca 180  
 aggaaggaaa gatgtgaata ggctgatggg caaaaaacca atttaccat cagttccagc 240  
 cttctctcaa gngaggcaa a 261

<210> 412  
 <211> 241  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(241)  
 <223> n = A,T,C or G

<400> 412  
 gttcaatgtt acctgacatt totacaacac cccactcacc gatgtattcg ttgccagtg 60  
 ggaacatacc agcctgaatt tggaaaaaat aattgtgttt cttgccagg aaatactacg 120  
 actgactttg atggctccac aaacataacc cagtgtaaaa acagaagatg tggaggggag 180  
 ctgggagatt tcaactgggta cattgaattc ccaaactacc cangcaatta cccagccaac 240  
 a 241

<210> 413  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(231)  
 <223> n = A,T,C or G

<400> 413  
 aactcttaca atccaagtga ctcatctgtg tgettgaatc ctttccactg tctcatctcc 60  
 ctcatccaag tttctagtac cttctctttg ttgtgaagga taatcaaact gaacaacaaa 120  
 aagtttactc tcctcatttg gaacctaaaa actctcttct tcctgggtct gagggctcca 180  
 agaatccttg aatcanttct cagatcattg gggacaccan atcaggaacc t 231

<210> 414  
 <211> 234  
 <212> DNA  
 <213> Homo sapiens



<400> 417  
nagtcttcag gcccatcagg gaagttcaca ctggagagaa gtcatacata tgtactgtat 60  
gtgggaaaagg ctttactctg agttcaaadc ttcaagccca tcagagagtc cacactggag 120



```

agaagccata caaatgcaat gagtgtggga agagcttcag gagggattcc cattatcaag 180
ttcatctagt ggtccacaca ggagagaaac cctataaatg tgagatatgt gggaagggt 240
tcantcaaag ttcgtatctt caaatccatc ngaaggncca cagtatanan aaacctttta 300
agt 303

```

<210> 418

<211> 328

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(328)

<223> n = A,T,C or G

<400> 418

```

tttttggcgg tgggtgggca gggacgggac angagtctca ctctgttgcc caggctggag 60
tgcacaggca tgatctcggc tcactacaac ccctgcctcc catgtccaag cgattcttgt 120
gcctcagcct tcctgttagc tagaattaca ggcacatgcc accacaccca gctagttttt 180
gtatttttag tagagacagg gtttcaccat gttggccagg ctggtctcaa actcctnacc 240
tcagnggtca ggctggtctc aaactcctga cctcaagtga tctgcccacc tcagcctccc 300
aaagtgtan gattacaggc cgtgagcc 328

```

<210> 419

<211> 389

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(389)

<223> n = A,T,C or G

<400> 419

```

cctcctcaag acggcctgtg gtccgcctcc cggcaaccaa gaagcctgca gtgccatattg 60
acccttgagc catggactgg agcctgaaag gcagcgtaca ccctgctcct gatcttgctg 120
cttgtttctt ctctgtggct ccattcatag cacagtgtgt gcaactgagga ttgtgcaggc 180
cgagcaaggc caagctggct caaagagcaa ccagtcaact ctgccacggg gtgccaggca 240
ccggttctcc agccaccaac ctactcgtct cccgcaaatt gcacatcagt tcttctaccc 300
taaaggtagg accaaagggc atctgctttt ctgaagtctt ctgctctatc agccatcacg 360
tggcagccac tcnggctgtg tcgacgcgg 389

```

<210> 420

<211> 408

<212> DNA

<213> Homo sapiens

<400> 420

```

gttctctcta actcctgcc gaaacagctc tctcaacat gagagctgca cccctcctcc 60
tggccagggc agcaagcctt agccttggtt tcttgtttct gctttttttc tggctagacc 120
gaagtgtact agccaaggag ttgaagtttg tgactttggt gtttcggcat ggagaccgaa 180

```



```
<210> 421
<211> 352
<212> DNA
<213> Homo sapiens
```

<400>	421						
gctcaaaaat	ctttttactg	atnggcacatg	ctacacaatc	attgactatt	acggaggcca	60	
gaggagaatg	aggcctggcc	tgggagccct	gtgcctacta	naagcacatt	agattatcca	120	
ttcactgaca	gaacagggtc	tttttggttc	cttcttctcc	accacnata	acttgacgtc	180	
ctccttcttg	aagattcttt	ggcagttgtc	tttgtcataa	cccacaggtg	tagaaacaag	240	
ggtgcaacat	gaaattttctg	tttcgtagca	agtgcacgtc	tcacaagttg	gcangtctgc	300	
cactccgagt	ttattgggtg	tttgtttcct	ttgagatcca	tgcatttctc	gg	352	

<400>	422						
atgccaccat	gctggcaatg	cagcggggcgg	tccaaggcct	gcataaccag	cccaagctgg	60	
cgatgatcga	cggcaaccgt	tgcccgaaagt	tgccgatgcc	agccgaagcg	gtggtcaagg	120	
gcgatagcaa	gggtgccggcg	atcgcgggcgg	cgtcaatcct	ggccaaggtc	agccgtgata	180	
gtgaaatggc	agctgtcgaa	ttgatctacc	cgggttatgg	catcggcggg	cataagggct	240	
atccgacacc	gggtgcacctg	gaagccttgc	agcggctggg	gccgacgccg	attcaccgac	300	
gcttcttcgg	ccggtacggc	tggcctatga	aaattat			337	

```
<220>
<221> misc_feature
<222> (1)...(310)
<223> n = A,T,C or G
```

<400> 423

gctcaaaaat	ctttttactg	atatggcatg	gctacacaat	cattgactat	tagaggccag	60
aggagaatga	ggcctggcct	gggagccctg	tgctactan	aagcncatta	gattatccat	120
tcactgacag	aacagggtctt	ttttgggtcc	ttcttctcca	ccacgatata	cttgcagtcc	180
tccttcttga	agattctttg	gcagttgtct	ttgtcataac	ccacaggtgt	anaaacaagg	240



```
<210> 424
<211> 370
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> (1)...(370)
<223> n = A,T,C or G
```

```
<210> 425
<211> 216
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> (1)...(216)
<223> n = A,T,C or G
```

```
<210> 426
<211> 596
<212> DNA
<213> Homo sapiens
```

<400>	426						
cttccagtga	ggataaccct	gttgccccg	gccgaggttc	tccattaggc	tctgattgat	60	
tggcagtcag	tgatggaagg	gtgtttctgat	cattccgact	gccccaaagg	tcgctggcca	120	
gctctctgtt	ttgctgagtt	ggcagtagga	cctaatttgt	taattaagag	tagatggtga	180	
gctgtccttg	tattttgatt	aacctaatgg	ccttcccagc	acgactcggg	ttcagctgga	240	
gacatcacgg	caacttttaa	tgaaatgatt	tgaagggccg	ttaagaggca	cttcccgtta	300	
ttaggcagtt	catctgcact	gataacttct	tggcagctga	gctggtcggg	gctgtggccc	360	
aaacgcacac	ttggcttttg	gttttgagat	acaactctta	atcttttagt	catgcttgag	420	



<212> DNA



$\langle 220 \rangle$  $\langle 222 \rangle \quad (1) \dots (507)$ 

<223> n = A, T, C or G

cttatcncaa	tggggctccc	aaacttggct	gtgcagtgga	aactccgggg	gaattttgaa	60
gaacactgac	acccatcttc	caccccgaca	ctctgattta	attgggctgc	agtgagaaca	120
gagcatcaat	ttaaaaagct	gcccagaatg	ttntcctggg	cagcgttgtg	atctttgccn	180
ccttcgtgac	tttatgcaat	gcacatgct	atttcatacc	taatgaggga	gttcaggag	240
attcaaccag	gatgtttcta	cncctgtggg	ttatgacaaa	gacaactgcc	aaagaatntt	300
caagaaggag	gactgcaagt	atatcgtggg	ggagaagaag	gacccaaaaa	agacctgttc	360
gtcagtgaa	tggataatct	aatgtgcttc	tagtaggcac	agggctcca	ggccaggcct	420
cattctctc	tggcctctaa	tagtcaatga	ttgtgtagcc	atgcctatca	gtaaaaagat	480
ttttgagcaa	aaaaaaaaaa	aaaaaaa				507

<211> 392

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

<222> (1) ... (392)

<223> n = A, T, C or G

<400> 431

gaaaattcag	aatggataaa	aacaaatgaa	gtacaaaata	tttcagattt	acatagcgat	60
aaacaagaaa	gcacttatca	ggaggactta	caaatggaag	tacactctan	aaccatcatc	120
tatcatggct	aaatgtgaga	ttagcacagc	tgtattattt	gtacattgca	aacacctaga	180
aagagatggg	aaacaaaatc	ccaggagttt	tgtgtgtgga	gtcctggggt	ttccaacaga	240
catcattcca	gcattctgag	attagggnga	ttggggatca	ttctggagtt	ggaatgttca	300
acaaaagtga	tgttgttagg	taaaatgtac	aacttctgga	tctatgcaga	cattgaaggt	360
gcaatqagtc	tggcttttac	tctgctgttt	ct			392

<210> 432

<211> 387

<212> DNA

<213> Homo sapiens

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (387)$ 

<223> n = A, T, C or G

<400> 432

```

ggatccnta   cataatcaaa   tatagctgta   gtacatgttt   tcattggngt   agattaccac   60
aaatgcaagg   caacatgtgt   agatctcttg   tcttattctt   ttgtctataa   tactgtattg   120
ngtagtccaa   gctctcggna   gtccagccac   tgnгааacat   gctcccttta   gattaacctc   180

```



```
<210> 433
<211> 281
<212> DNA
<213> Homo sapiens
```

```
<400> 433
ttcaactagc anagaanact gcttcagggg gtgtaaaaatg aaaggcttcc acgcagttat 60
ctgattaaag aacactaaga gagggacaag gctagaagcc gcaggatgtc tacactatag 120
caggcnctat ttgggttggc tggaggagct gtggaaaaca tggagagatt ggcgctggag 180
atcgccgtgg ctattctctn ttgntattac accagngagg ntctctgtnt gcccaactgg 240
tnnaaaaccg ntatacaata atgatagaat aggacacaca t                               281
```

<400>	434						
ttttaaaata	agcatttagt	gctcagtc	ccc	tactgagta	c	tctttctctc	ccctcctctg 60
aattttaattc	tttcaacttg	caatttgcaa	ggattacaca	tttcaactgtg	atgtatatattg		120
tgttgcaaaaa	aaaaaaaaagt	gtctttgttt	aaaattactt	ggtttgtgaa	tccatcttgc		180
tttttccccca	ttggaactag	tcattaaccc	atctctgaac	tggtagaaaa	acatctgaag		240
agctagtctta	tcagcatctg	acaggtgaat	tggatggttc	tcagaaccat	ttcaaccaga		300
cagcctgtttt	ctatcctggt	taataaatta	gtttgggttc	tctacatgca	taacaaaccc		360
tgctccaatc	tgtcacataa	aagtctgtga	cttgaagttt	agtcagcacc	cccaccaaac		420
tttattttttc	tatgtgtttt	ttgcaacata	tgagtgtttt	gaaaataaag	taccatgtc		480
ttta							484

<400>	435						
gcgccgctca	gagcagggtca	otttctgcct	tccacgtcct	ccttcaagga	agcccatgt	60	
gggtagcttt	caatatcgca	ggttcttact	cctctgcctc	tataagctca	aaccaccaa	120	
cgatcgggca	agtaaacccc	ctccctcgcc	gacttcggaa	ctggcgagag	ttcagcgcag	180	
atgggcctgt	ggggaggggg	caagatagat	gagggggagc	ggcatggtgc	ggggtgacct	240	
cttgagagaga	ggaaaaaggc	cacaagaggg	gctgccaccg	ccactaacgg	agatggccct	300	
ggtagagacc	tttgggggtc	tggaaacctct	ggactcccca	tgctctaact	cccacactct	360	
gctatcagaa	acttaaacct	gaggattttc	tctgtttttc	actcgcaata	aattcagagc	420	



424

<211> 667

<212> DNA

<213> Homo sapiens

 $\langle 220 \rangle$ 

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (667)$ 

<223> n = A,T,C or G

<400> 436

accttgggaa	nactctcaca	atataaaggg	tctgtagactt	tactccaaat	tccaaaaagg	60
tccttggccat	gtaatctctga	aagttttccc	aaggtagcta	taaaatcctt	ataagggtgc	120
agcctcttct	ggaattcctc	tgatttcaaa	gtctcactct	caagttcttg	aaaacgaggg	180
cagttctctga	aaggcaggta	tagcaactga	tcttcagaaa	gaggaaactgt	gtgcaccggg	240
atgggctgcc	agagtaggat	aggattccag	atgctgacac	cttctggggg	aaacagggct	300
gccaggtttg	tcatagcact	catcaaagtc	cggtcacgt	ctgtgcttcg	aatataaacc	360
tgttcatgtt	tataggactc	attcaagaat	tttctatatc	tctttcttat	atactctcca	420
agttcataat	gctgctccat	gccagctgg	gtgagttggc	caaatccttg	tggccatgag	480
gattccttta	tggggtcagt	gggaaaggty	tcaatgggac	ttcggctctcc	atgccgaaac	540
accaaagtca	caaacttcaa	ctccttggt	agtacacttc	ggtctagcca	gaaaaaaagc	600
agaaacaaga	agccaaggct	aaggcttgct	gcctgccag	gaggaggggg	gcagctctca	660
tgttgag						667

<210> 437

<211> 693

<212> DNA

<213> Homo sapiens

<400> 437

ctacgtctca	accctcattt	ttaggtaagg	aatcttaagt	ccaaagatat	taagtgactc	60
acacagccag	gtaaggaaag	ctggattggc	acactaggac	tctaccatac	cgggttttgt	120
taaagctcag	gttaggaggc	tgataagctt	ggaaggaaact	tcagacagct	ttttcagatc	180
ataaaaagata	attcttagcc	catgttcttc	tccagagcag	acctgaaatg	acagcacagc	240
aggtaactcct	ctattttcac	ccctcttgct	tctactctct	ggcagtcaga	cctgtgggag	300
gccatgggag	aaagcagctc	tctggatggt	tgtacagatc	atggactatt	ctctgtggac	360
catttctcca	ggttacccta	ggtgtcacta	ttggggggac	agccagcatc	tttagctttc	420
atttgagttt	ctgtctgtct	tcagtagagg	aaacttttgc	tcttcacact	tcacatctga	480
acacctaaact	gctgttgctc	ctgagggtgg	gaaagacaga	tatagagctt	acagtattta	540
tcctattttct	aggcaactgag	ggctgtgggg	taccttgtgg	tgccaaaaca	gatcctgttt	600
taaggacatg	ttgcttcaga	gatgtctgta	actatctggg	ggctctgttg	gctctttacc	660
ctgcacatg	tgctctcttg	gctgaaaatg	acc			693

<210> 438

<211> 360

<212> DNA

<213> Homo sapiens

<400> 438



```

ctgcttatca caatgaatgt tctcctgggc agcgttgtga tctttgccac ctctgtgact 60
ttatgcaatg catcatgcta tttcatacct aatgagggag ttccaggaga ttcaaccagg 120
atgtttctac acctgtgggt tatgacaaag acaactgcc aagaatcttc aagaaggagg 180
actgcaagta tatctggtgg agaagaagga cccaaaaaag acctgttctg tcagtgaatg 240
gataatctaa tgtgcttcta gtaggcacag ggctcccagg ccaggcctca ttctcctctg 300
gcctctaata gtcaataatt gtgtagccat gcctatcagt aaaaagattt ttgagcaaac 360

```

<210> 439

<211> 431

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(431)

<223> n = A,T,C or G

<400> 439

```

gttcctnnta actcctgcc aaaaacagctc tcctcaacat gagagctgca cccctcctcc 60
tggccagggc agcaagcctt agccttggct tcttgtttct gctttttttc tggctagacc 120
gaagtgtact agccaaggag ttgaagtttg tgactttggg gtttcggcat ggagaccgaa 180
gtcccattga cacctttccc actgacccca taaaggaatc ctcatggcca caaggatttg 240
gccaaactcac ccagctgggc atggagcagc attatgaact tggagagtat ataagaaaga 300
gatatagaaa attcttgaat gagtcctata aacatgaaca ggtttatatt cgaagcacag 360
acgttgaccg gactttgatg agtgctatga caaacctggc agcccgtcga cgcggccgcg 420
aatttagtag t
431

```

<210> 440

<211> 523

<212> DNA

<213> Homo sapiens

<400> 440

```

agagataaag cttaggtcaa agttcataga gttcccatga actatatgac tggccacaca 60
ggatcttttg tatttaagga ttctgagatt ttgcttgagc aggattagat aaggctgttc 120
tttaaagtgc tgaaatggaa cagatttcaa aaaaaaaccc cacaatctag ggtgggaaca 180
aggaaggaaa gatgtgaata ggctgatggg caaaaaacca atttacccat cagttccagc 240
cttctctcaa ggagaggcaa agaaaggaga tacagtggag acatctggaa agttttctcc 300
actggaaaac tgctactatc tgtttttata tttctgttaa aatatatgag gctacagaac 360
taaaaattaa aacctctttg tgtcccttgg tcctggaaca tttatgttcc ttttaaagaa 420
acaaaaatca aactttacag aaagatttga tgtatgtaat acatatagca gctcttgaag 480
tatatatatc atagcaaata agtcatctga tgagaacaag cta
523

```

<210> 441

<211> 430

<212> DNA

<213> Homo sapiens

<400> 441

```

gttcctccta actcctgcc aaaaacagctc tcctcaacat gagagctgca cccctcctcc 60
tggccagggc agcaagcctt agccttggct tcttgtttct gctttttttc tggctagacc 120

```



```

gaagtgtact agccaaggag ttgaagtttg tgacttttggg gtttcggcat ggagaccgaa 180
gtcccattga cacctttccc actgacccca taaaggaatc ctcatggcca caaggatttg 240
gccaaactcac ccagctgggc atggagcagc attatgaact tggagagtat ataagaaaga 300
gatatagaaa attccttgaat gagtcctata aacatgaaca ggtttatatt cgaagcacag 360
acgttgaccg gactttgatg agtgctatga caaacctggc agcccgtcga cgcggccgcg 420
aatttagtag                                     430

```

<210> 442

<211> 362

<212> DNA

<213> Homo sapiens

<400> 442

```

ctaaggaatt agtagtgctt ccatcacttg tttggagtgt gctattctaa aagattttga 60
tttcctggaa tgacaattat attttaactt tgggtggggga aagagttata ggaccacagt 120
cttcacttct gatacttgta aattaatctt ttattgcact tgttttgacc attaagctat 180
atgttttagaa atgggtcattt tacggaaaaa ttagaaaaat tctgataata gtgcagaata 240
aatgaattaa tgttttactt aattttatatt gaactgtcaa tgacaaataa aaattctttt 300
tgattatttt ttgttttcat ttaccagaat aaaaactaag aattaaaagt ttgattacag 360
tc                                             362

```

<210> 443

<211> 624

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1) ... (624)

<223> n = A,T,C or G

<400> 443

```

tttttttttt gcaacacaat atacatcaca gtgaaatgtg taatccttgc aaattgcaag 60
ttgaaagaat taaattcaga ggaggggaga gaaagagtac tcagtaggga ctgagcacta 120
aatgcttatt ttaaaagaaa tgtaaagagc agaaagcaat tcaggctacc ctgccttttg 180
tgctggctag tactccggtc ggtgtcagca gcacgtggca ttgaacattg caatgtggag 240
cccaaaccac agaaaatggg gtgaaattgg ccaactttct attaaacttg cttcctgttt 300
tataaaatat tgtgaataat atcacctact tcaaagggca gttatgagga ttaaataaac 360
taacgcctac aaaacactta aacatagata acatagggtgc aagtactatg tatctggtac 420
atggtaaaca tccttattat taaagtcaac gctaaaatga atgtgtgtgc atatgctaata 480
agtacagaga gagggcactt aaaccaacta agggcctgga ggggaaggttt cctggaaaga 540
ngatgcttgt gctgggtcca aatcttggtc tactatgacc ttggccaaat tattttaaact 600
ttgtccctat ctgctaaaca gatc                                     624

```

<210> 444

<211> 425

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature



<223> n = A, T, C or G

gacacatcatt	nntcttgc	tctttgagaa	taagaagatc	agtaaatagt	tcagaagtgg	60
gaagctttgt	ccaggcctgt	gtgtgaaccc	aatgttttgc	ttagaaatag	aacaagtaag	120
ttcattgcta	tagcataaca	caaaatttgc	ataagtgggt	gtcagcaaat	ccttgaatgc	180
tgcttaatgt	gagaggttgg	taaaatcctt	tgtgcaacac	tctaactccc	tgaatgtttt	240
gctgtgctgg	gacctgtgca	tgccagacaa	ggccaagctg	gctgaaagag	caaccagcca	300
cctctgcaat	ctgccacctc	ctgctggcag	gatttgtttt	tgcctcctgt	gaagagccaa	360
ggaggcacca	gggcataagt	gagtagactt	atggtcgacg	cggccgcgaa	tttagtagta	420
qtaga						425

<213> Homo sapiens

<223> n = A,T,C or G

catgtttatg	nttttggtatt	actttgggca	cctagtgttt	ctaaatcgtc	tatcattctt	60
ttctgttttt	caaaagcaga	gatggccaga	gtctcaacaa	actgtatctt	caagtctttg	120
tgaattcttt	tgcattgtggc	agattattgg	atgtagtctt	ctttaactag	catataaatc	180
tgggtgtgtt	cagataaatg	aacagcaaaa	tgtggtggaa	ttaccatttg	gaacattgtg	240
aatgaaaaat	tgtgtctcta	gattatgtaa	caaataacta	tttcctaacc	attgatcttt	300
ggatttttat	aatcctactc	acaaatgact	aggcttctcc	tcttgatttt	tgaagcagtg	360
tgggtgctgg	attgataaaa	aaaaaaaaag	tgcagcgggc	cgcgaattta	gtag	414

<213> Homo sapiens

<223> n = A, T, C or G

acaaattaga	anaaagtgcc	agagaacacc	acataccttg	tccggaacat	tacaatggct	60
tctgcatgca	tgggaagtgt	gagcattcta	tcaatatgca	ggagccatct	tgcagggtgtg	120
atgctggtta	tactggacaa	cactgtgaaa	aaaaggacta	cagtgttcta	tacgtttgttc	180
ccggtcctgt	acgatttcag	tatgtcttaa	tgcgagctgt	gattggaaca	attcagattg	240
ctgtcatctg	tgtggtggtc	ctctgcatca	caagggccaa	actttaggta	atagcattgg	300
actgagattt	gtaaaccttc	caaccttcca	ggaaatgcc	cagaagcaac	agaattcaca	360
gacagaagca	aaatacaggg	cactacagtt	cagacaatac	aacaagagcg	tccacgaggt	420
taatctaaag	ggagcatgtt	tcacaagtgc	tggactaccg	agagcttgga	ctacacaata	480



```

cagtattata gacaaaagaa taagacaaga gatctacaca tgttgccctg catttgtggg 540
aatctacacc aatgaaaaca tgtactacag ctatatattga ttatgtatgg atatatattga 600
aatagtatac attgtcttga tgttttttct g 631

```

<210> 447

<211> 585

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(585)

<223> n = A,T,C or G

<400> 447

```

ccttgaggaaa antntcacaa tataaagggt cgtagacttt actccaaatt ccaaaaagggt 60
cctggccatg taatcctgaa agttttccca aggtagctat aaaatcctta taagggtgca 120
gcctcttctg gaattcctct gattttcaaag tctcactctc aagttcttga aaacgagggc 180
agttcctgaa aggcaggtat agcaactgat cttcagaaag aggaactgtg tgcaccggga 240
tggtctgcca gagtaggata ggattccaga tgctgacacc ttctggggga aacagggctg 300
ccaggtttgt catagcactc atcaaagtcc ggtcaacgtc tgtgcttcga atataaacct 360
gttcattgttt ataggactca ttcaagaatt ttctatatct ctttcttata tactctccaa 420
gttcataatg ctgctccatg cccagctggg tgagttggcc aaatccttgt ggccatgagg 480
attcctttat ggggtcagtg ggaaagggtg caatgggaact tcggtctcca tgccgaaaca 540
ccaaagtcac aaacttcaac tccttggtta gtacacttcg gtcta 585

```

<210> 448

<211> 93

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(93)

<223> n = A,T,C or G

<400> 448

```

tgctcgtggg tcattctgan nnccgaactg accntgccag ccttgccgan gggccnccat 60
ggctccctag tgccctggag agganggggc tag 93

```

<210> 449

<211> 706

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(706)

<223> n = A,T,C or G

<400> 449



```

ccaagttcat gctntgtgct ggacgctgga cagggggcaa aagcnnttgc tegtgggtca 60
ttctgancac cgaactgacc atgccagccc tgccgatggt cctccatggc tccctagtgc 120
cctggagagg aggtgtctag tcagagagta gtccctggaag gtggcctctg ngaggagcca 180
cggggacagc atcctgcaga tggtcgggcg cgccccattc gccattcagg ctgcgcaact 240
gttgggaagg gcgatcggtg cgggcctctt cgctattacg ccagctggcg aaagggggat 300
gtgctgcaag gcgattaaat tgggtaacgc caggggttttc ccagtcncga cgttgtaaaa 360
cgacggccag tgaattgaat ttaggtgacn ctatagaaga gctatgacgt cgcatgcacg 420
cgtacgtaag cttggatcct ctagagcggc cgccactac tactaaattc gcggccgcgt 480
cgacgtggga tcncactga gagagtggag agtgacatgt gctggacnct gtccatgaag 540
cactgagcag aagctggagg cacaacgcnc cagacactca cagctactca ggaggctgag 600
aacaggttga acctgggagg tggagggttc aatgagctga gatcaggccn ctgcncccca 660
gcatggatga cagagtgaaa ctccatctta aaaaaaaaaa aaaaaa 706

```

<210> 450

<211> 493

<212> DNA

<213> Homo sapiens

<400> 450

```

gagacggagt gtcactctgt tgcccaggct ggagtgcagc aagacactgt ctaagaaaaa 60
acagttttta aaggtaaaaa aacataaaaa gaaatatcct atagtggaaa taagagagtc 120
aaatgaggct gagaacttta caaagggatc ttacagacat gtgcgcaata tcaactgcatg 180
agcctaagta taagaacaac ctttggggag aaaccatcat ttgacagtga ggtacaattc 240
caagtcagggt agtgaaatgg gtggaattaa actcaaatta atcctgccag ctgaaacgca 300
agagacactg tcagagagtt aaaaagttag ttctatccat gaggtgattc cacagtcttc 360
tcaagtcaac acatctgtga actcacagac caagttctta aaccactgtt caaactctgc 420
tacacatcag aatcacctgg agagctttac aaactcccat tgccgagggg cgacgcggcc 480
gcgaatttag tag 493

```

<210> 451

<211> 501

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(501)

<223> n = A,T,C or G

<400> 451

```

gggcgcgtcc cattcgccat tcaggctgcg caactgttgg gaagggcgat cgggtgcgggc 60
ctcttcgcta ttacgccagc tggcgaaaagg gggatgtgct gcaagggcat taagttgggt 120
aacgccaggg ttttccagtc cncgacgttg taaaacgacg gccagtgaat tgaatttagg 180
tgacnctata gaagagctat gacgtcgcat gcacgcgtac gtaagcttgg atcctctaga 240
gcggccgcct actactacta aattcgcggc cgcgtcgacg tgggatacnc actgagagag 300
tgagagtgta catgtgctgg acnctgtcca tgaagcactg agcagaagct ggaggcacia 360
cgcncagac actcacagct actcaggagg ctgagaacag gttgaacctg ggagggtggag 420
gttgcaatga gctgagatca ggccnctgcn cccagcatg gatgacagag tgaaactcca 480
tcttaaaaaa aaaaaaaaaa a 501

```

<210> 452



<211> 51  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(51)  
 <223> n = A,T,C or G

<400> 452  
 agacggtttc accntttacaa cnccttttag gatgggnntt ggggagcaag c 51

<210> 453  
 <211> 317  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(317)  
 <223> n = A,T,C or G

<400> 453  
 tacatcttgc tttttcccca ttggaactag tcattaaccc atctctgaac tggtagaaaa 60  
 acatctgaag agctagtcta tcagcatctg gcaagtgaat tggatggttc tcagaaccat 120  
 ttcacccana cagcctgttt ctatcctgtt taataaatta gtttgggttc tctacatgca 180  
 taacaaaccc tgctccaate tgtcacataa aagtctgtga cttgaagttt antcagcacc 240  
 cccaccaaac tttatttttc tatgtgtttt ttgcaacata tgagtgtttt gaaaataagg 300  
 taccatgtc tttatta 317

<210> 454  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 454  
 ttcgaggtag aatcaactct cagagtgtag tttccttcta tagatgagtc agcattaata 60  
 taagccacgc cagctcttg aaggagtctt gaattctcct ctgctcactc agtagaacca 120  
 agaagaccaa attcttctgc atcccagctt gcaaacaaaa ttgttcttct aggtctccac 180  
 ccttcctttt tcagtgttcc aaagctcctc acaatttcat gaacaacagc t 231

<210> 455  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 455  
 taccaaagag ggcataataa tcagtctcac agtagggttc accatcctcc aagtgaaaaa 60  
 cattgttccg aatgggcttt ccacaggcta cacacacaaa acaggaaaca tgccaagttt 120  
 gtttcaacgc attgatgact tctccaagga tcttcctttg gcatcgacca cattcagggg 180  
 caaagaattt ctcatagcac agctcacaat acagggtctc tttctcctct a 231



<210> 456  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 456  
 ttggcaggta cccttacaaa gaagacacca taccttatgc gttattaggt ggaataatca 60  
 ttccattcag tattatcggt attattcttg gagaaaccct gtctgtttac tgtaaccctt 120  
 tgcactcaaa ttctttatc aggaataact acatagccac tatttacaaa gccattggaa 180  
 cctttttatt tgggtgcagct gctagtcagt ccctgactga cattgccaag t 231

<210> 457  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(231)  
 <223> n = A,T,C or G

<400> 457  
 cgaggtagcc aggggtctga aaatctctnn ttantagtc gatagcaaaa ttgttcatca 60  
 gcattcctta atatgatctt gctataatta gatttttctc cattagagtt catacagttt 120  
 tatttgattt tattagcaat ctctttcaga agacccttga gatcattaag ctttgtatcc 180  
 agttgtctaa atcgatgcct catttctctt gaggtgtcgc tggcttttgt g 231

<210> 458  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 458  
 aggtctgggt ccccccaatt ccactccctt ctactctctc taggactggg ctgggccaag 60  
 agaagagggg tggttaggga agcgttgag acctgaagcc ccaccctcta ccttccttca 120  
 acaccctaac cttgggtaac agcatttgga attatcattt gggatgagta gaatttccaa 180  
 ggtcctgggt taggcatttt gggggggccag accccaggag aagaagattc t 231

<210> 459  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 459  
 ggtaccgagg ctgcgtgaca cagagaaacc ccaacgcgag gaaaggaatg gccagccaca 60  
 ccttcgcgaa acctgtggtg gccaccagt cctaaccgga caggacagag agacagagca 120  
 gccctgcact gttttccctc caccacagcc atcctgtccc tcattggctc tgtgctttcc 180  
 actatacaca gtcaccgtcc caatgagaaa caagaaggag caccctccac a 231

<210> 460



```
<400> 464
gtactctaag attttatcta agttgccttt tctgggtggg aaagtttaac cttagtgact 60
aaqqacatca catatgaaga atgtttaagt tggaggtggc aacgtgaatt gcaaacaggg 120
```



cctgcttcag tgactgtgtg cctgtagtcc cagctactcg ggagtctgtg tgaggccagg 180  
 ggtgccagcg caccagctag atgctctgta acttctaggg cccattttcc c 231

<210> 465  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 465  
 catgttggtg tagctgtggt aatgctggct gcattctcaga cagggttaac ttcagctcct 60  
 gtggcacaatt agcaacaaat tctgacatca tatttatggg ttctgtatct ttgttgatga 120  
 aggatggcac aattttttgct tgtgttcata atatactcag attagttcag ctccatcaga 180  
 taaactggag acatgcagga cattagggta gtgtttagc tctggtaatg a 231

<210> 466  
 <211> 231  
 <212> DNA  
 <213> Homo sapiens

<400> 466  
 caggtacctc tttccattgg ataactgtgct agcaagcatg ctctccgggg tttttttaat 60  
 ggccttcgaa cagaacttgc cacataccca ggtataatag tttctaacat ttgccagga 120  
 cctgtgcaat caaatattgt ggagaattcc ctagctggag aagtcacaaa gactataggc 180  
 aataatggag accagtccca caagatgaca accagtcgtt gtgtgcggt g 231

<210> 467  
 <211> 311  
 <212> DNA  
 <213> Homo sapiens

<400> 467  
 gtacaccctg gcacagtcca atctgaactg gttcggcact catctttcat gagatggatg 60  
 tgggtggcttt tctccttttt catcaagact cctcagcagg gagcccagac cagcctgcac 120  
 tgtgccttaa cagaaggctc tgagattcta agtgggaatc atttcagtga ctgtcatgtg 180  
 gcatgggtct ctgcccaagc tcgtaatgag actatagcaa ggcggctgtg ggacgtcagt 240  
 tgtgacctgc tgggcctccc aatagactaa caggcagtgc cagttggacc caagagaaga 300  
 ctgcagcaga c 311

<210> 468  
 <211> 3112  
 <212> DNA  
 <213> Homo sapiens

<400> 468  
 cattgtgttg ggagaaaaac agaggggaga tttgtgtggc tgcagccgag ggagaccagg 60  
 aagatctgca tgggtgggaag gacctgatga tacagagttt gataggagac aattaaaggc 120  
 tggaaggcac tggatgctg atgatgaagt ggactttcaa actggggcac tactgaaacg 180  
 atgggatggc cagagacaca ggagatgagt tggagcaagc tcaataacaa agtgggttcaa 240  
 cgaggacttg gaattgcatg gagctggagc tgaagttag cccaattgtt tactagttag 300  
 gtgaatgtgg atgattggat gatcatttct catctctgag cctcaggttc cccatccata 360  
 aaatgggata cacagtatga tctataaagt gggatatagt atgatctact tcaactgggtt 420



```

atttgaagga tgaattgaga taatttattt caggtgccta gaacaatgcc cagattagta 480
catttggtgg aactgagaaa tggcataaca ccaaatttaa tatatgtcag atgttactat 540
gattatcatt caatctcata gttttgtcat ggcccaattt atcctcactt gtgcctcaac 600
aaattgaact gttaacaaag gaatctctgg tcctgggtaa tggctgagca ccactgagca 660
tttccattcc agttggcttc ttgggtttgc tagctgcac actagtcac ttaaataaat 720
gaagttttaa catttctcca gtgatttttt tatctcacct ttgaagatac tatgttatgt 780
gattaaataa agaacttgag aagaacaggt ttcattaaac ataaaatcaa tgtagacgca 840
aattttctgg atgggcaata cttatgttca caggaaatgc tttaaaatat gcagaagata 900
attaaatggc aatggacaaa gtgaaaaact tagacttttt tttttttttt ggaagtatct 960
ggatgttcct tagtcactta aaggagaact gaaaaatagc agtgagttcc acataatcca 1020
acctgtgaga ttaaggctct ttgtggggaa ggacaaagat ctgtaaattt acagtttcct 1080
tccaaagcca acgtcgaatt ttgaaacata tcaaagctct tcttcaagac aaataatcta 1140
tagtacatct ttcttatggg atgcacttat gaaaaatggt ggctgtcaac atctagtcac 1200
tttagctctc aaaatggttc attttaagag aaagttttag aatctcatat ttattcctgt 1260
ggaaggacag cattgtggct tggactttat aaggctctta ttcaactaaa taggtgagaa 1320
ataagaaaag ctgctgactt taccatctga ggccacacat ctgctgaaat ggagataatt 1380
aacatcacta gaaacagcaa gatgacaata taatgtctaa gtagtgacat gtttttgcac 1440
atttccagcc cctttaaata tccacacaca caggaagcac aaaaggaagc acagagatcc 1500
ctgggagaaa tgcccgcccg ccatcttggg tcatcgatga gcctcgccct gtgcctggtc 1560
ccgcttgtga gggaaggaca ttagaaaatg aattgatgtg ttccttaaag gatgggcagg 1620
aaaacagatc ctgttgtgga tatttatattg aacgggatta cagatttgaa atgaagtcac 1680
aaagtgagca ttaccaatga gagggaaaaca gacgagaaaa tcttgatggc ttcacaagac 1740
atgcaacaaa caaaatggaa tactgtgatg acatgaggca gccaaagctgg ggaggagata 1800
accacggggc agagggtcag gattctggcc ctgctgccta aactgtgcgt tcataaccaa 1860
atcatttcat atttctaacc ctcaaaacaa agctgttgta atatctgac tctacggttc 1920
cttctgggcc caacattctc catatatcca gccacactca tttttaatat ttagttccca 1980
gatctgtact gtgaccttcc tacactgtag aataacatta ctcattttgt tcaaagacc 2040
ttegtgttgc tgcctaatat gtagctgact gtttttccca aggagtgttc tggcccagg 2100
gatctgtgaa caggctggga agcatctcaa gatctttcca ggggttatac tactagcaca 2160
cagcatgatc attacggagt gaattatcta atcaacatca tcctcagtgt ctttgcccat 2220
actgaaattc atttccact tttgtgcccc ttctcaagac ctcaaaatgt cattccatta 2280
atatcacagg attaaacttt ttttttaacc tggagaat caatgttaca tgcagctatg 2340
ggaatttaac tacatatttt gttttccagt gcaaagatga ctaagtcctt tatccctccc 2400
ctttgtttga ttttttttcc agtataaagt taaaatgctt agccttgtag tgaggctgta 2460
tacagccaca gcctctcccc atccctccag ccttatctgt catcaccatc aaccctccc 2520
atgcacctaa acaaaatcta acttgtaatt ccttgaacat gtcaggcata cattattcct 2580
tctgcctgag aagctcttcc ttgtctctta aatctagaat gatgtaaagt tttgaataag 2640
ttgactatct tacttcatgc aaagaaggga cacatatgag attcatcatc acatgagaca 2700
gcaaatacta aaagtgtaat ttgattataa gagtttagat aaatatatga aatgcaagag 2760
ccacagaggg aatgtttatg gggcacgttt gtaagcctgg gatgtgaagc aaaggcagg 2820
aacctcatag tatcttatat aatatacttc atttctctat ctctatcaca atatccaaca 2880
agcttttcac agaattcatg cagtgcaaat ccccaaagg aacctttatc catttcatgg 2940
tgagtgcgct ttagaatttt ggcaaatcat actggtcact tatctcaact ttgagatgtg 3000
tttgtccttg tagttaattg aaagaaatag ggcactcttg tgagccactt tagggttcac 3060
tcctggcaat aaagaattta caaagagcaa aaaaaaaaaa aaaaaaaaaa aa 3112

```

<210> 469

<211> 2229

<212> DNA

<213> Homo sapiens



&lt;400&gt; 469

```

agctctttgt aaattcttta ttgccaggag tgaaccctaa agtggctcac aagagtgtcc 60
tatttctttc aattaactac aaggacaaac acatctcaaa gttgagataa gtgaccagta 120
tgatttgcca aaattctaaa gcgcactcac catgaaatgg ataaaggtta cctttgggga 180
tttgcaactgc atgaattctg tgaaaagctt gttggatatt gtgatagaga tagagaaatg 240
aagtatatta tataagatac tatgaggttc cctgcctttg cttcacatcc caggcttaca 300
aacgtgtccc ataaacattc cctctgtggc tcttgcatth catatattha tctaaactct 360
tataatcaaa tacactttta gtatttgctg tctcatgtga tgatgaatct catatgtgtc 420
ccttctttgc atgaagtaag atagtcaact tattcaaaac tttacatcat tctagattta 480
agagacaagg aagagcttct caggcagaag gaataatgta tgctgacat gttcaaggaa 540
ttacaagtta gattttgttt aggtgcatgg gaggggttga tgggtgatgac agataaggct 600
ggaggggatgg ggagaggctg tggctgtata cagcctcagt acaaggctaa gcattttaac 660
tttatactgg aaaaaaatc aaacaaaggg gagggataaa ggacttagtc atctttgcac 720
tggaaaacaa aatatgtaat taaattccca tagctgcatg taacattgaa ttctccagg 780
ttaaaaaaaa agttaatcct gtgatattaa tggaaatgaca ttttgaggct ttgagaatgg 840
gcacaaaagt gggaaatgaa tttcagtatg ggcaaagaca ctgaggatga tgttgattag 900
ataattcact ccgtaatgat catgctgtgt gctagtaagt ataaccctgg aaagatcttg 960
agatgcttcc cagcctgttc acagatcccc tgggccagaa cactccttag gaaaaacagt 1020
cagctacata ttaggcagca acacgaaggg tctttgaaca aaatgagtaa tgttattcta 1080
cagtgtagaa aggtcacagt acagatctgg gaactaaata ttaaaaatga gtgtggctgg 1140
atatatggag aatgttgggc ccagaaggaa ccgtagagat cagatattac aacagctttg 1200
ttttgagggt tagaaatatg aaatgatttg gttatgaacg cacagtttag gcagcagggc 1260
cagaatcctg accctctgcc ccgtggttat ctcctccca gcttggctgc ctcatgtcat 1320
cacagtattc cattttgttt gttgcatgtc ttgtgaagcc atcaagattt tctcgtctgt 1380
tttctctca ttggaatgc tcactttgtg acttcatttc aaatctgtaa tcccgttcaa 1440
ataaatatcc acaacaggat ctgttttctt gccatcctt taaggaacac atcaattcat 1500
tttctaattg ccttccctca caagcgggac caggcacagg gcgaggctca tcatgaccc 1560
aagatggcgg ccgggcattt ctcccaggga tctctgtgct tctttttgtg cttcctgtgt 1620
gtgtggatat ttaagggggc tggaaatgtg caaaaacatg tcaactacta gacattatat 1680
tgtcatcttg ctgtttctag tgatgttaat tatctccatt tcagcagatg tgtggcctca 1740
gatggtaaag tcagcagcct ttcttatttc tcacctggaa atacatacga ccatttgagg 1800
agacaaatgg caaggtgtca gcataccctg aacttgagtt gagagctaca cacaatatta 1860
ttggtttccg agcatcacaa acaccctctc tgtttcttca ctgggcacag aattttaata 1920
cttatttcag tgggctgttg gcaggaacaa atgaagcaat ctacataaag tcaactagtgc 1980
agtgcctgac acacaccatt ctcttgaggt cccctctaga gatccacag gtcatatgac 2040
ttcttgggga gcagtggctc acacctgtaa tcccagcact ttgggaggct gaggcagggtg 2100
ggtcacctga ggtcaggagt tcaagaccag cctggccaat atggtgaaac cccatctcta 2160
ctaaaaatac aaaaattagc tgggcgtgct ggtgcatgcc tgtaatccca gcccacac 2220
aatggaatt 2229

```

&lt;210&gt; 470

&lt;211&gt; 2426

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 470

```

gtaaatcttt tattgccagg agtgaaccct aaagtggctc acaagagtgc cctatttctt 60
tcaattaact acaaggacaa acacatctca aagttgagat aagtgaccag tatgatttgc 120
caaaattcta aagcgactc accatgaaat ggataaagg tacccttggg gatttgcact 180
gcatgaattc tgtgaaaagc ttgttgata ttgtgataga gatagagaaa tgaagtatat 240
tatataagat actatgaggt tccctgcctt tgcttcacat cccaggctta caaacgtgcc 300

```



```
<210> 471
<211> 812
<212> DNA
<213> Homo sapiens
```

<400> 471						
gaacaaaatg	agtaatgtta	ttctacagtg	tagaaaaggtc	acagtacaga	tctgggaact	60
aaatattaaa	aatgagtgtg	gctggatata	tggagaatgt	tgggcccaga	aggaaccgta	120
gagatcagat	attacaacag	ctttgttttg	agggttagaa	atatgaaatg	atttggttat	180
gaacgcacag	tttaggcagc	agggccagaa	tctgaccct	ctgccccgtg	gttatctcct	240
cccagccttg	gctgcctcat	gtcatcacag	tattccattt	tgtttgttgc	atgtcttggtg	300
aagccatcaa	gattttctcg	tctgttttcc	tctcattggt	aatgctcact	ttgtgacttc	360
atttcaaatc	tgtaatcccg	ttcaataaaa	tatccacaac	aggatctggt	ttcctgcccc	420
tcctttaaag	aacacatcaa	ttcattttct	aatgtccttc	cctcacaagc	gggaccaggc	480



```
<210> 472
<211> 515
<212> DNA
<213> Homo sapiens
```

<400> 472						
acggagactt	attttctgat	attgtctgca	tatgtatggt	tttaagagtc	tggaaatagt	60
cttatgactt	tcctatcatg	cttattaata	aataatacag	cccagagaag	atgaaaatgg	120
gttcagaat	tattggtcct	tgcagcccgg	tgaatctcag	caagaggaac	caccaactga	180
caatcaggat	attgaacctg	gacaagagag	agaaggaaca	cctccgatcg	aagaacgtaa	240
agtagaaggt	gattgccagg	aaatggatct	ggaaaagact	cggagtgagc	gtggagatgg	300
ctctgatgta	aaagagaaga	ctccacctaa	tcctaagcat	gctaagacta	aagaagcagg	360
agatgggcag	ccataagtta	aaaagaagac	aagctgaagc	tacacacatg	gctgatgtca	420
cattgaaaat	gtgactgaaa	atttgaaaat	tctctcaata	aagtttgagt	tttctctgaa	480
gaaaaaaaaa	naaaaaaaaa	aaanaaaaaa	aaaaa			515

<400> 473						
cgcatgccgg	ggaagcccaa	gctggctcga	agagccacca	gccacctgtg	caaggggtggg	60
cctggaccag	ttggaccagc	caccaagctc	acctactcaa	ggaagcaggg	atggccaggt	120
tgcaacagcc	tgagtggctg	ccacctgata	gctgatggag	cagaggcctg	aggaaaatca	180
gatggcacat	ttagctcttt	aatggatctt	aagttaattt	ttctataaag	catatggcac	240
cagtccatgc	ctcagagctc	gtatggcact	gcggaccaca	gcaggccgag	ttcccaggat	300
tgccatccag	gggggccttc	tgtagccctg	gccagacctt	gcagaggttg	ctgggtgctc	360
tttgagcgag	ctcggcctcc	ctggcatgca	caggccccag	gtactgacac	gctgctctga	420
gtgagcttgt	cctgccttgg	ctgccacctt	actgctgatg	gagcagcggc	cttaggaaaa	480
gcaaattggcg	ctgtagccca	actttagggt	agaagaagat	gtaccatgtc	cggccgctag	540
ttggtgactg	gtgcacctgc	tcctggcgta	cccttgccaga	gggtgggtgg	tgctcttttg	600
ccagcttggc	cttgcttggc	atgcacaagc	ctcagtgcaa	caactgtcct	acaaatggag	660
acacagagag	gaaacaagca	gcgggctcag	gagcaggggtg	tgtgctgcct	ttggggctcc	720
agtccatgcc	tcgggtcgta	tggtactgca	ggcttctttg	ttgccaaagag	gcggaccaca	780
ggccttcttg	aggaggactt	tacgttcaag	tgcagaaagc	agccaaaatt	accatccatg	840
agactaagcc	ttctgtggcc	ctggcgagac	ttaaaatttg	tgccaaggca	ggacaagctc	900
actcgagca	gcgtgtcagt	agctggggcc	tatgcatgcc	gggcagggcc	gggctggctg	960
aaqgaqcaac	cagccacctc	tgcaaggggtg	gcctagtgc	aggcggagca	tccaccacct	1020







```
<210> 474
<211> 1594
<212> DNA
<213> Homo sapiens
```

<400>	474						
atttatggat	cattaatgcc	tcttttagtag	tttagagaaa	acgtcaaaag	aaatggcccc	60	
agaataagct	tcttgatttg	taaaattcta	tgtcattggc	tcaaatttgt	atagtatctc	120	
aaaatataaa	tatatagaca	tctcagataa	tatatattgaa	atagcaaatt	cctgttagaa	180	
aataatagta	cttaactaga	tgagaataac	aggtcgccat	tatttgaatt	gtctcctatt	240	
cgtttttcat	ttgtttgtgt	actcatgttt	tacttatgag	ggatatatat	aacttccact	300	
gttttcagaa	ttattgtatg	cagtcagtat	gagaatgcaa	tttaagtttc	cttgatgctt	360	
tttcacactt	ctattactag	aaataagaat	acagtaatat	tggcaaagaa	aattgaccag	420	
ttcaataaaa	tttttttagta	aatctgattg	aaaataaaca	ttgcttatgg	ctttcttaca	480	
tcaatattgt	tatgtcctag	acaccttatc	tgaaattacg	gcttcaaaat	tctaattatg	540	
tgcaaattgt	taaaatatca	atactttatg	ttcaagctgg	ggcctcttca	ggcgtcctgg	600	
gctgagagag	aaagatgcta	gctccgcaag	ccggagaggg	aacaccgcca	cattgtttaca	660	
cggacacacc	gccacgtgga	cacatgacca	gactcacatg	tacagacaca	cggagacatt	720	
accacatgga	gacaccgtca	cacagtcaca	cggacacact	ggcatagtca	catggacgga	780	
cacacagaca	tatggagaaa	tcacatggac	acaccaccac	actatcacag	ggacacagac	840	



```

acacggagac atcaccacat ggacacactg tcacactacc acagggacac gagacatcac 900
actgtcacat ggacacacca tcacacacat gaacacaccg acacactgcc atatggacac 960
tggcacacac actgccacac tgtcacatgg acacacctcc acaccatcac accaccacac 1020
acactgcctg tggacacaag gacacacaga cactgtcaca cagatacaca aaacactgtc 1080
acacggagac atcaccatgc agatacacca ccactctggt gccgtctgaa ttaccctgct 1140
ggggggacag cagtggcata ctcatgccta agtgactggc tttcacccca gtagtgattg 1200
ccctccatca acactgcccc ccccagggtg gggctacccc agcccatctt tacaaaaacag 1260
ggcaagggtga actaatggag tgggtggagg agttggaaga aatcccagcg tcagtcaccg 1320
ggatagaatt cccaaggaac cctctttttg gaggatggtt tccatttctg gaggcgatct 1380
gccgacaggg tgaatgcctt cttgcttgct ttctggggaa tcagagagag tccgttttgt 1440
ggtgggaaga gtgtggctgt gtactttgaa ctctgtaaa ttctctgact catgtccaca 1500
aaaccaacag ttttgtgaat gtgtctggag gcaagggaag ggccactcag gatctatgtt 1560
gaaggggaaga ggctgggggc tggagtattc gctt 1594

```

<210> 475

<211> 2414

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (33)

<223> n=A,T,C or G

<400> 475

```

cccaacacaa tggctttata agaatgcttc acntgtgaaa aacaaatata aaagtcttct 60
ttagatttat ttttaaggac aaatctttat tccatgttta atttatttag ctttccctgt 120
agctaataat tcatgctgaa cacattttaa atgctgtaaa ttagataat gtaatttatg 180
tatcattaat gcctctttag tagtttagag aaaacgtcaa aagaaatggc ccagaataa 240
gcttcttgat ttgtaaaatt ctatgtcatt ggctcaaatt tgtatagtat ctcaaaatat 300
aaatatatag acatctcaga taatatattt gaaatagcaa attcctgtta gaaaataata 360
gtacttaact agatgagaat aacaggctgc cattatttga attgtctcct attcgttttt 420
catttggtgt gttactcatg ttttacttat ggggggatat atataacttc cgctgttttc 480
agaagtattg tatgcagtca gtatgagaat gcaatttaag tttccttgat gctttttcac 540
acttctatta ctgaaaataa gaatacagta atattggcaa agaaaattga ccagttcaat 600
aaaatttttt agtaaatctg attgaaaata aacattgctt atggctttct tacatcaata 660
ttgttatgtc ctgacacact tatctgaaat tacggcttca aaattctaata tatgtgcaaa 720
tgtgtaaaat atcaatactt tatgttcaag ctggggcctc ttcaggcgctc ctgggctgag 780
agagaaagat gctagctccg caagccgggg agggaaacacc gccacattgt tacatggaca 840
caccgccacg tggacacatg accagactca catgtacaga cacacggaga cattaccaca 900
tggagacacc gtcacacagt cacacgagca cactggcata gtcacatgga cggacacaca 960
gacatatgga gaaatcacac tgacacacca ccacactatc acagggacac agacacacgg 1020
agacatcacc acatggacac actgtcacac taccacaggg acacgagaca tcacactgtc 1080
acatggacac accatcacac acatgaacac accgacacac tgccatatgg aactgtccac 1140
acacactgcc aactgtcac atggacacac ctccatacca tcacaccacc acacacactg 1200
ccatgtggac acaaggacac acagacactg tcacacagat acacaaaaca ctgtcacacg 1260
gagacatcac catgcagata caccaccaca tggacatagc accagacact ctgccacaca 1320
gatacaccac cacacagaaa tgcggacaca ctgccacaca gacaccacca catcgttgcc 1380
acactttcat gtgtcagctg gcggtgtggg cccacgact ctgggctcta atcgagaaat 1440
tacttgagca tatagtgaag gcaaaaattt tttttatttt ctgggtaacc aagcgcgact 1500
ctgtctcaaa aaaagaaaaa aaaagcaata tactgtgtaa tcgttgacag cataattcac 1560

```



```

tattatgtag atcggagagc agaggattct gaatgcatga acatatcatt aacatttcaa 1620
tacattactc ataattactg atgaactaaa gagaaaccaa gaaattatgg tgatagttat 1680
attgacctgg agaaatgtag acacaaaaga accgtaagat gagaaatgtg ttaacacagt 1740
ctataagggc atgcaagaat aaaaataggg gagaaaacag gagagttttt caagagcttt 1800
ctggtcattg aagtcaactt gtatcgggta attttttaaa ggtttattta catgcaataa 1860
actgcacata cttcaattgt acattttggg aattcttggc atttgtagct ctataaaacc 1920
agcaacatat taaaatagca aacatatcca ttacctttac caccaaaagt ttcttgtgtt 1980
ttttctactc actttttcct gcctatcccc ccctctcttc cacaggtaac cactgatcca 2040
cttccagtca ctatccatga gtttttattt ccaaatacat gaaatcatat gaattttctg 2100
tttttcctgt tggagcccaa ggagcaaggg cagaatgagg aacatgatgt ttcttwccga 2160
cagtactca tgacgtctcc atccaggact gaggggggca tccttctcca tctaggactg 2220
ggggcatcct tctccatcca gtattggggg tcatccttct ccaccagta ttgggggtca 2280
tcctcctcca tccaggacct gaggggtgtc cttttctgcg cttccttggg tggcagttct 2340
tcccttcattg tttatagtra cttaccatta aatcactgtg ccgttttttc ctaaaaataa 2400
aaaaaaaaaa aaaa 2414

```

<210> 476

<211> 3434

<212> DNA

<213> Homo sapiens

<400> 476

```

ctgtgctgca aatggggcca tatagaggaa aggagcagct ggctctggag catggtgtgc 60
actccctttg ggccttcagt ccattgtctc tgggtcgtat gacactgcgg gcttgttggg 120
tgccaagagg cagaccacag gtcattctga ggaggacttt atgttccagt ccagaaagca 180
gccagtggta ccaccagggg gacttgtgct tctgtggccc aggccagacg tagaatttga 240
caaagtcagg acggtctcag tcagagcagc atgtcgggct ccggggcctg tgcattgccg 300
gcagggccag gctggcttaa ggagcaagca gccacctctg ttaggggtgt gcctggagca 360
ggtggagcag ccaccaacct cacgcactga aagaagcagg gatggccagg ttccaacatc 420
ctgagtggct gccacctgat ggctgatgga gcagaggcct gaggaaaagc agatggcact 480
gctttgtagt gctgttcttt gtctctcttg atctttttca gttaatgtct gttttatcag 540
agactaggat tgcaaacctt gctctttttt gctttccatt tgcttggtta atattcctcc 600
atccctttat tttaaacctc tgtgtgtctt tgcacatgag atgggtctcc tgaatacagg 660
acaacaatgg gtctttactc tttatccaac ttgccagtct gtgtctttta actggggcat 720
ttagcccatc tacatttaag tttagtattt gttacatgtg aaatttatcc tgtcatgatg 780
ttgctagctt tttatttttc ccattagttt gcagtttctt tatagtgtca atgggtctta 840
caattcgata tgtttttgta gtggctggta ctgggttttc ctttctacgt ttagtgtctc 900
cttcaggagc tcttgtaaca caagaatgtg gatttatctc ttgtaaggta aatatgtgga 960
tttattctgg gactgtattc tatggccttt accccaagaa tcattacttt ttaaaatgca 1020
attcaaatta gcataaaaca tttacagcct atggaaaggc ttgtggcatt agaatcctta 1080
tttataggat tattttgtgt ttttttgaga tatggctctt gtcactgagg cagaagtgcc 1140
gtggtttgat cataattcac cacagccctg aactcttgag tccaagccat cttttgcct 1200
taatctccca accagttgga tctacaagca taaggcatca tgcgtggcta attttttcac 1260
gttttttttt tttttgtcga gattatggta tcaactgtgt gctctggctg atctcaaag 1320
tttgacctca agggatcttt ctgccacagc ctccataaagt gctaggatta tatgcatgat 1380
acaccatgcc tattgtagag tattacatta ttttcaaagt cttattgtta gagccattta 1440
ttgccttttg cctaaataac tcaatataat atctctgaaa cttttttttg acaaattttg 1500
gggcgtgatg atgagagaag ggggtttgaa acttttctaa aagagttaac ttagagccat 1560
ttaagaaagg aaaaaacaca aattatcaga aaaacaacag taagatcaag tgcaaaagt 1620
ctgtggcaaa gatgatgaga gtaaagaata tatgtttgtg actcatgggt gcttttactt 1680
tgttcttgaa tttctgagta cgggttaaca tttaaagaat ctacattata gataacattt 1740

```







```
<210> 479
<211> 223
<212> PRT
<213> Homo sapiens
```



Met Tyr Arg His Thr Glu Thr Leu Pro His Gly Asp Thr Val Thr Gln  
5 10 15

Gly Glu Ile Thr Leu Thr His His His Thr Ile Thr Gly Thr Gln Thr  
35 40 45

Arg Asp Ile Thr Leu Ser His Gly His Thr Ile Thr His Met Asn Thr  
65 70 75 80

His Gly His Thr Ser Ile Pro Ser His His His Thr His Cys His Val  
100 105 110

Thr Arg Arg His His His Ala Asp Thr Pro Pro His Gly His Ser Thr  
130 135 140

Cys His Thr Asp Thr Thr Thr Ser Leu Pro His Phe His Val Ser Ala  
165 170 175

Thr Tyr Ser Glu Gly Lys Ile Phe Phe Tyr Phe Leu Gly Asn Gln Ala  
195 200 205

<210> 480

<211> 145

<212> PRT

<213> Homo sapiens

<400> 480



Ala Asp Gly Pro Trp Pro Tyr Leu Phe Val Arg Arg Thr Asp Val Pro  
65 70 75 80



Trp Met Ala Met Phe Pro Gln Pro Glu Trp Leu Pro Pro Asp Gly



140

```
<400> 483
Met Glu Thr Gln Arg Gly Asn Lys Gln Arg Ala Gln Glu Gln Gly Val
      5                               10                      15

Cys Cys Leu Trp Gly Ser Ser Pro Cys Leu Gly Ser Tyr Gly Thr Ala
      20                                25                      30

Gly Phe Leu Val Ala Lys Arg Arg Thr Thr Gly Leu Leu Glu Glu Asp
      35                                  40                      45

Phe Thr Phe Lys Cys Arg Lys Gln Pro Lys Leu Pro Ser Met Arg Leu
      50                                  55                      60

Ser Leu Leu Trp Pro Trp Arg Asp Leu Lys Phe Val Pro Arg Gln Asp
      65                                  70                      75                      80

Lys Leu Thr Arg Ser Ser Val Ser Val Ala Gly Ala Tyr Ala Cys Arg
      85                                  90                      95

Ala Gly Pro Gly Trp Leu Lys Glu Gln Pro Ala Thr Ser Ala Arg Val
     100                                105                      110

Arg Leu Val Gln Ala Glu His Pro Pro Pro His Pro Leu Glu Glu Val
     115                                120                      125

Gly Met Ala Arg Phe Pro Gln Pro Glu Cys Leu Pro Pro Tyr Cys
     130                                135                      140
```

<400> 484															
Thr	Ala	Ala	Ser	Asp	Asn	Phe	Gln	Leu	Ser	Gln	Gly	Gly	Gln	Gly	Phe
1				5					10					15	
Ala	Ile	Pro	Ile	Gly	Gln	Ala	Met	Ala	Ile	Ala	Gly	Gln	Ile		
			20					25					30		

```
<210> 485
<211> 31
<212> DNA
<213> Artificial Sequence
```



&lt;220&gt;

&lt;223&gt; Made in a lab

&lt;400&gt; 485

gggaagctta tcacctatgt gccgcctctg c

31

&lt;210&gt; 486

&lt;211&gt; 27

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Made in a lab

&lt;400&gt; 486

gcgaattctc acgctgagta tttggcc

27

&lt;210&gt; 487

&lt;211&gt; 36

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Made in a lab

&lt;400&gt; 487

cccgaattct tagctgccca tccgaacgcc ttcac

36

&lt;210&gt; 488

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Made in a lab

&lt;400&gt; 488

gggaagcttc ttccccggct gcaccagctg tgc

33

&lt;210&gt; 489

&lt;211&gt; 19

&lt;212&gt; PRT

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Made in a lab

&lt;400&gt; 489

Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala

1

5

10

15

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000  
1001  
1002  
1003  
1004  
1005  
1006  
1007  
1008  
1009  
1010  
1011  
1012  
1013  
1014  
1015  
1016  
1017  
1018  
1019  
1020  
1021  
1022  
1023  
1024  
1025  
1026  
1027  
1028  
1029  
1030  
1031  
1032  
1033  
1034  
1035  
1036  
1037  
1038  
1039  
1040  
1041  
1042  
1043  
1044  
1045  
1046  
1047  
1048  
1049  
1050  
1051  
1052  
1053  
1054  
1055  
1056  
1057  
1058  
1059  
1060  
1061  
1062  
1063  
1064  
1065  
1066  
1067  
1068  
1069  
1070  
1071  
1072  
1073  
1074  
1075  
1076  
1077  
1078  
1079  
1080  
1081  
1082  
1083  
1084  
1085  
1086  
1087  
1088  
1089  
1090  
1091  
1092  
1093  
1094  
1095  
1096  
1097  
1098  
1099  
1100  
1101  
1102  
1103  
1104  
1105  
1106  
1107  
1108  
1109  
1110  
1111  
1112  
1113  
1114  
1115  
1116  
1117  
1118  
1119  
1120  
1121  
1122  
1123  
1124  
1125  
1126  
1127  
1128  
1129  
1130  
1131  
1132  
1133  
1134  
1135  
1136  
1137  
1138  
1139  
1140  
1141  
1142  
1143  
1144  
1145  
1146  
1147  
1148  
1149  
1150  
1151  
1152  
1153  
1154  
1155  
1156  
1157  
1158  
1159  
1160  
1161  
1162  
1163  
1164  
1165  
1166  
1167  
1168  
1169  
1170  
1171  
1172  
1173  
1174  
1175  
1176  
1177  
1178  
1179  
1180  
1181  
1182  
1183  
1184  
1185  
1186  
1187  
1188  
1189  
1190  
1191  
1192  
1193  
1194  
1195  
1196  
1197  
1198  
1199  
1200  
1201  
1202  
1203  
1204  
1205  
1206  
1207  
1208  
1209  
1210  
1211  
1212  
1213  
1214  
1215  
1216  
1217  
1218  
1219  
1220  
1221  
1222  
1223  
1224  
1225  
1226  
1227  
1228  
1229  
1230  
1231  
1232  
1233  
1234  
1235  
1236  
1237  
1238  
1239  
1240  
1241  
1242  
1243  
1244  
1245  
1246  
1247  
1248  
1249  
1250  
1251  
1252  
1253  
1254  
1255  
1256  
1257  
1258  
1259  
1260  
1261  
1262  
1263  
1264  
1265  
1266  
1267  
1268  
1269  
1270  
1271  
1272  
1273  
1274  
1275  
1276  
1277  
1278  
1279  
1280  
1281  
1282  
1283  
1284  
1285  
1286  
1287  
1288  
1289  
1290  
1291  
1292  
1293  
1294  
1295  
1296  
1297  
1298  
1299  
1300  
1301  
1302  
1303  
1304  
1305  
1306  
1307  
1308  
1309  
1310  
1311  
1312  
1313  
1314  
1315  
1316  
1317  
1318  
1319  
1320  
1321  
1322  
1323  
1324  
1325  
1326  
1327  
1328  
1329  
1330  
1331  
1332  
1333  
1334  
1335  
1336  
1337  
1338  
1339  
1340  
1341  
1342  
1343  
1344  
1345  
1346  
1347  
1348  
1349  
1350  
1351  
1352  
1353  
1354  
1355  
1356  
1357  
1358  
1359  
1360  
1361  
1362  
1363  
1364  
1365  
1366  
1367  
1368  
1369  
1370  
1371  
1372  
1373  
1374  
1375  
1376  
1377  
1378  
1379  
1380  
1381  
1382  
1383  
1384  
1385  
1386  
1387  
1388  
1389  
1390  
1391  
1392  
1393  
1394  
1395  
1396  
1397  
1398  
1399  
1400  
1401  
1402  
1403  
1404  
1405  
1406  
1407  
1408  
1409  
1410  
1411  
1412  
1413  
1414  
1415  
1416  
1417  
1418  
1419  
1420  
1421  
1422  
1423  
1424  
1425  
1426  
1427  
1428  
1429  
1430  
1431  
1432  
1433  
1434  
1435  
1436  
1437  
1438  
1439  
1440  
1441  
1442  
1443  
1444  
1445  
1446  
1447  
1448  
1449  
1450  
1451  
1452  
1453  
1454  
1455  
1456  
1457  
1458  
1459  
1460  
1461  
1462  
1463  
1464  
1465  
1466  
1467  
1468  
1469  
1470  
1471  
1472  
1473  
1474  
1475  
1476  
1477  
1478  
1479  
1480  
1481  
1482  
1483  
1484  
1485  
1486  
1487  
1488  
1489  
1490  
1491  
1492  
1493  
1494  
1495  
1496  
1497  
1498  
1499  
1500  
1501  
1502  
1503  
1504  
1505  
1506  
1507  
1508  
1509  
1510  
1511  
1512  
1513  
1514  
1515  
1516  
1517  
1518  
1519  
1520  
1521  
1522  
1523  
1524  
1525  
1526  
1527  
1528  
1529  
1530  
1531  
1532  
1533  
1534  
1535  
1536  
1537  
1538  
1539  
1540  
1541  
1542  
1543  
1544  
1545  
1546  
1547  
1548  
1549  
1550  
1551  
1552  
1553  
1554  
1555  
1556  
1557  
1558  
1559  
1560  
1561  
1562  
1563  
1564  
1565  
1566  
1567  
1568  
1569  
1570  
1571  
1572  
1573  
1574  
1575  
1576  
1577  
1578  
1579  
1580  
1581  
1582  
1583  
1584  
1585  
1586  
1587  
1588  
1589  
1590  
1591  
1592  
1593  
1594  
1595  
1596  
1597  
1598  
1599  
1600  
1601  
1602  
1603  
1604  
1605  
1606  
1607  
1608  
1609  
1610  
1611  
1612  
1613  
1614  
1615  
1616  
1617  
1618  
1619  
1620  
1621  
1622  
1623  
1624  
1625  
1626  
1627  
1628  
1629  
1630  
1631  
1632  
1633  
1634  
1635  
1636  
1637  
1638  
1639  
1640  
1641  
1642  
1643  
1644  
1645  
1646  
1647  
1648  
1649  
1650  
1651  
1652  
1653  
1654  
1655  
1656  
1657  
1658  
1659  
1660  
1661  
1662  
1663  
1664  
1665  
1666  
1667  
1668  
1669  
1670  
1671  
1672  
1673  
1674  
1675  
1676  
1677  
1678  
1679  
1680  
1681  
1682  
1683  
1684  
1685  
1686  
1687  
1688  
1689  
1690  
1691  
1692  
1693  
1694  
1695  
1696  
1697  
1698  
1699  
1700  
1701  
1702  
1703  
1704  
1705  
1706  
1707  
1708  
1709  
1710  
1711  
1712  
1713  
1714  
1715  
1716  
1717  
1718  
1719  
1720  
1721  
1722  
1723  
1724  
1725  
1726  
1727  
1728  
1729  
1730  
1731  
1732  
1733  
1734  
1735  
1736  
1737  
1738  
1739  
1740  
1741  
1742  
1743  
1744  
1745  
1746  
1747  
1748  
1749  
1750  
1751  
1752  
1753  
1754  
1755  
1756  
1757  
1758  
1759  
1760  
1761  
1762  
1763  
1764  
1765  
1766  
1767  
1768  
1769  
1770  
1771  
1772  
1773  
1774  
1775  
1776  
1777  
1778  
1779  
1780  
1781  
1782  
1783  
1784  
1785  
1786  
1787  
1788  
1789  
1790  
1791  
1792  
1793  
1794  
1795  
1796  
1797  
1798  
1799  
1800  
1801  
1802  
1803  
1804  
1805  
1806  
1807  
1808  
1809  
1810  
1811  
1812  
1813  
1814  
1815  
1816  
1817  
1818  
1819  
1820  
1821  
1822  
1823  
1824  
1825  
1826  
1827  
1828  
1829  
1830  
1831  
1832  
1833  
1834  
1835  
1836  
1837  
1838  
1839  
1840  
1841  
1842  
1843  
1844  
1845  
1846  
1847  
1848  
1849  
1850  
1851  
1852  
1853  
1854  
1855  
1856  
1857  
1858  
1859  
1860  
1861  
1862  
1863  
1864  
1865  
1866  
1867  
1868  
1869  
1870  
1871  
1872  
1873  
1874  
1875  
1876  
1877  
1878  
1879  
1880  
1881  
1882  
1883  
1884  
1885  
1886  
1887  
1888  
1889  
1890  
1891  
1892  
1893  
1894  
1895  
1896  
1897  
1898  
1899  
1900  
1901  
1902  
1903  
1904  
1905  
1906  
1907  
1908  
1909  
1910  
1911  
1912  
1913  
1914  
1915  
1916  
1917  
1918  
1919  
1920  
1921  
1922  
1923  
1924  
1925  
1926  
1927  
1928  
1929  
1930  
1931  
1932  
1933  
1934  
1935  
1936  
1937  
1938  
1939  
1940  
1941  
1942  
1943  
1944  
1945  
1946  
1947  
1948  
1949  
1950  
1951  
1952  
1953



```

<210> 490
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Made in a lab

<400> 490
Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys
 1             5             10             15
Leu Ser His Ser
                20

<210> 491
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Made in a lab

<400> 491
Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
 1             5             10             15
Thr Gly Phe Thr
                20

<210> 492
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> Made in a lab

<400> 492
Ala Leu Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr
 1             5             10             15
Leu Ala Ser Leu
                20

<210> 493
<211> 20
<212> PRT
<213> Artificial Sequence

<220>

```



<400> 493

<210> 494

<212> PRT

<220>

<400> 494

<210> 495

<212> PRT

 $\langle 220 \rangle$ 

<400> 495

<210> 496

<211> 21

<212> PRT

 $\langle 220 \rangle$ 

<400> 496

Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu Leu  
1 5 10 15  
Pro Pro Pro Pro Ala  
20



<220>  
<223> Made in a lab

```
<210> 498
<211> 20
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

```
<210> 499
<211> 20
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

```
<210> 500
<211> 20
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab



```
<210> 501
<211> 20
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

```
<210> 502
<211> 414
<212> DNA
<213> Homo Sapien
```

<400> 502						
caccatgggag	acaggcctgc	gctggcctttt	cctggctcgc	gtgctcaaag	gtgtccaatg	60
tcagtcggtg	gaggagtcog	ggggctgcct	ggtcacgcct	gggacacctt	tgacantcac	120
ctgtagagtt	tttgaatng	acctcagtag	caatgcaatg	agctgggtcc	gccaggctcc	180
agggaaaggg	ctggaatgga	tggagccat	tgataattgt	ccacantacg	cgacctgggc	240
gaaaggccga	tnatnatntt	ccaaaacctn	gaccacgggtg	gatttgaaaa	tgaccagtcc	300
gacaaccgag	gacacggcca	cctattttttg	tggcagaatg	aatactggta	atagtggttg	360
gaagaatat	tggggccag	gcacctggt	caccgtntcc	tcagggcaac	ctaa	414

```
<210> 503
<211> 379
<212> DNA
<213> Homo Sapien
```

<400>	503					
atnCGatggT	gcttggTcaa	aggtgtccag	tgtcagtcg	tggaggagtc	cgggggTcgc	60
ctggTcacgc	ctgggacacc	cctgacactc	acctgcaccg	tntctggatt	ngacatcagt	120
agctatggag	tgagctgggt	ccgccaggct	ccaggggaagg	ggctgggnata	catcggatca	180
ttagtagtag	tggtacattt	tacgcgagct	gggcgaaagg	ccgattccacc	atttccaaaa	240
cctngaccac	ggtggatttg	aaaatcacca	gtttgacaac	cgaggacacg	gccacctatt	300
tntgtgccag	aggggggttt	aattataaag	acatttgggg	cccaggccacc	ctggTcaccg	360
tntccttagg	gcaacctaa					379

```
<210> 504
<211> 19
<212> PRT
<213> Artificial Sequence
```



<223> Made in a lab

Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp Ser Pro Tyr Phe Lys Glu  
1 5 10 15  
Asn Ser Ala

<213> Artificial Sequence

<223> Made in a lab

Lys Glu Asn Ser Ala Phe Pro Pro Phe Cys Cys Asn Asp Asn Val Thr  
1 5 10 15  
Asn Thr Ala Asn  
20

<213> Homo Sapien

atggagacag	gctcgcgctg	gcttctcctg	gtcgcgtgcg	tcaaaggtgt	ccagtgtcag	60
tcgctggagg	agtcoggggg	tcgcctggtc	acgcctggga	caccctgac	actcacctgc	120
accgtctctg	gatttcacct	cagtagcaat	gcaatgatct	gggtccgcca	ggctccaggg	180
aaggggctgg	aatacatcgg	atacattagt	tatggtggtg	gcgcatacta	cgcgagctgg	240
gtgaaaggcc	gattcaccat	ctccaaaacc	tcgaccacgg	tggatctgag	aatgaccagt	300
ctgacaaccg	aggacacggc	cacctatttc	tgtgccagaa	atagtgattt	tagtggtatg	360
ttgtggggcc	caggcacccct	ggtcaccgtc	tcctcagggc	aacctaa		407

<213> Homo Sapien

atggagacag	gcctgcgctg	gcttctcctg	gtcgctgtgc	tcaaaggtgt	ccagtgtcag	60
tcggtggagg	agtccggggg	tcgcctggtc	acgcctggga	caccctgac	actcacctgt	120
acagtctctg	gattctccct	cagcaactac	gacctgaact	gggtccgcca	ggctccaggg	180
aaggggctgg	aatggatcgg	gatcattaat	tatgttggtg	ggacggacta	cgcgaactgg	240
gcaaaaggcc	ggttcaccat	ctccaaaacc	tcgaccaccg	tggatctcaa	gatcgccagt	300
ccqacaaccg	aggacacggc	cacctatttc	tgtgccagag	ggtggaagtg	cgatgagtct	360



```

ggtcctgtgt tgcgcattctg gggcccaggc accctgggtca cgtctcctt agggcaacct 420
aa 422

```

```

<210> 508
<211> 411
<212> DNA
<213> Homo Sapien

```

```

<400> 508
atggagacag gcctcgctgg cttctcctgg tcgtgtgtgt caaagggtgt cagtgtcagt 60
cggtggagga gtccgggggt cgctgggtca cgctggggac acccctgaca ctcacctgca 120
cagtctcttg aatcgacctc agtagctact gcatgagctg ggtccgccag gctccaggga 180
aggggctgga atggatcgga atcattggta ctctgggtga cacatactac gcgaggtggg 240
cgaaaaggccg attcaccatc tccaaaacct cgaccacggg gcatntgaaa atcnccagtc 300
cgacaaccga ggacacggcc acctatttct gtgccagaga tcttcgggat ggtagtagta 360
ctggttatta taaaatctgg ggcccaggca ccttgggtcac cgtctccttg g 411

```

```

<210> 509
<211> 15
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Made in a lab

```

```

<400> 509
Leu Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
1           5           10          15

```

```

<210> 510
<211> 15
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Made in a lab

```

```

<400> 510
Pro Glu Tyr Asn Arg Pro Leu Leu Ala Asn Asp Leu Met Leu Ile
1           5           10          15

```

```

<210> 511
<211> 15
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Made in a lab

```

```

<400> 511

```

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



```
<210> 512
<211> 15
<212> PRT
<213> Artificial Sequence
```

<400> 512

```
<210> 513
<211> 15
<212> PRT
<213> Artificial Sequence
```

<400> 513

```
<210> 514
<211> 15
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

<400> 514

```
<210> 515
<211> 15
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

<400> 515

Met Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg  
1 5 10 15



```
<220>  
<223> Made in a lab
```

```
<210> 517
<211> 15
<212> PRT
<213> Artificial Sequence
```

```
<210> 518
<211> 15
<212> PRT
<213> Artificial Sequence
```

```
<210> 519
<211> 17
<212> PRT
<213> Artificial Sequence
```

<210> 520



<220>  
<223> Made in a lab

```
<210> 521
<211> 21
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

```
<210> 522
<211> 20
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

```
<210> 523
<211> 254
<212> PRT
<213> Artificial Sequence
```

<220>  
<223> Made in a lab

<400> 523  
Met Ala Thr Ala Gly Asn Pro Trp Gly Trp Phe Leu Gly Tyr Leu Ile



1	5	10	15
Leu Gly Val Ala Gly Ser Leu Val Ser Gly Ser Cys Ser Gln Ile Ile			
20	25	30	
Asn Gly Glu Asp Cys Ser Pro His Ser Gln Pro Trp Gln Ala Ala Leu			
35	40	45	
Val Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln			
50	55	60	
Trp Val Leu Ser Ala Thr His Cys Phe Gln Asn Ser Tyr Thr Ile Gly			
65	70	75	80
Leu Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met			
85	90	95	
Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu			
100	105	110	
Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu			
115	120	125	
Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala			
130	135	140	
Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg			
145	150	155	160
Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu Glu			
165	170	175	
Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys			
180	185	190	
Ala Gly Gly Gly Gln Xaa Gln Xaa Asp Ser Cys Asn Gly Asp Ser Gly			
195	200	205	
Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly			
210	215	220	
Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn Leu			
225	230	235	240
Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser			
245	250		

&lt;210&gt; 524

&lt;211&gt; 765

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 524

atggccacag	caggaaatcc	ctggggctgg	ttcttgggggt	acctcattcct	tggtgtcgca	60
ggatcgctcg	tctctggtag	ctgcagccaa	atcataaacg	gcgaggactg	cagcccgcac	120
tcgcagccct	ggcaggcggc	actggtcatg	gaaaacgaat	tggtctgctc	gggcgtcctg	180
gtgcatccgc	agtgggtgct	gtcagccgca	cactgtttcc	agaactccta	caccatcggg	240
ctgggcctgc	acagtcttga	ggccgaccaa	gagccagggg	gccagatggg	ggaggccagc	300
ctctccgtac	ggcaccacaga	gtacaacaga	cccttgctcg	ctaacgacct	catgctcatc	360
aagttggacg	aatccgtgtc	cgagtctgac	accatccgga	gcattcagcat	tgcttcgcag	420
tgccctaccg	cggggaaactc	ttgcctcggt	tctggctggg	gtctgctggc	gaacggcaga	480
atgcctaccg	tgctgcagtg	cgtgaacgtg	tcgggtggtg	ctgaggaggt	ctgcagtaag	540
ctctatgacc	cgtgtgacca	ccccagcatg	ttctgcgccg	gctggaggga	agaccagaag	600
gactcctgca	acggtgactc	tggggggccc	ctgatctgca	acgggtactt	gcagggcctt	660
gtgtctttcg	gaaaagcccc	gtgtggccaa	gttggcgtgc	caggtgtcta	caccaacctc	720
tgcaaatcca	ctgagtggat	agagaaaacc	gtccaggcca	gttaa		765



<210> 525  
 <211> 254  
 <212> PRT  
 <213> Homo sapien

<400> 525  
 Met Ala Thr Ala Gly Asn Pro Trp Gly Trp Phe Leu Gly Tyr Leu Ile  
 1 5 10 15  
 Leu Gly Val Ala Gly Ser Leu Val Ser Gly Ser Cys Ser Gln Ile Ile  
 20 25 30  
 Asn Gly Glu Asp Cys Ser Pro His Ser Gln Pro Trp Gln Ala Ala Leu  
 35 40 45  
 Val Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln  
 50 55 60  
 Trp Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly  
 65 70 75 80  
 Leu Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met  
 85 90 95  
 Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu  
 100 105 110  
 Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu  
 115 120 125  
 Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala  
 130 135 140  
 Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg  
 145 150 155 160  
 Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu Glu  
 165 170 175  
 Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys  
 180 185 190  
 Ala Gly Gly Gly Gln Asp Gln Lys Asp Ser Cys Asn Gly Asp Ser Gly  
 195 200 205  
 Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly  
 210 215 220  
 Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn Leu  
 225 230 235 240  
 Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser  
 245 250

<210> 526  
 <211> 963  
 <212> DNA  
 <213> Homo sapiens

<400> 526  
 atgagttcct gcaacttcac acatgccacc tttgtgctta ttggtatccc aggattagag 60  
 aaagcccatt tctgggttgg cttccccctc ctttccatgt atgtagtggc aatgttttga 120  
 aactgcatcg tgggtcttcat cgtaaggacg gaacgcagcc tgcacgctcc gatgtacctc 180  
 tttctctgca tgcttgacgc cattgaacctg gccttatcca catccaccat gcctaagatc 240  
 cttgcccttt tctgggtttga ttcccgagag attagctttg aggctgtct taccagatg 300



<210> 527

&lt;212&gt; PRT

<213> Homo sapiens

Met Ser Ser Cys Asn Phe Thr His Ala Thr Phe Val Leu Ile Gly Ile  
5 10 15

Pro Gly Leu Glu Lys Ala His Phe Trp Val Gly Phe Pro Leu Leu Ser  
20 25 30

Met Tyr Val Val Ala Met Phe Gly Asn Cys Ile Val Val Phe Ile Val  
35 40 45

Arg Thr Glu Arg Ser Leu His Ala Pro Met Tyr Leu Phe Leu Cys Met  
50 55 60

Leu Ala Ala Ile Asp Leu Ala Leu Ser Thr Ser Thr Met Pro Lys Ile  
65 70 75 80

Leu Ala Leu Phe Trp Phe Asp Ser Arg Glu Ile Ser Phe Glu Ala Cys  
85 90 95

Leu Thr Gln Met Phe Phe Ile His Ala Leu Ser Ala Ile Glu Ser Thr  
100 105 110

Ile Leu Leu Ala Met Ala Phe Asp Arg Tyr Val Ala Ile Cys His Pro  
115 120 125

Leu Arg His Ala Ala Val Leu Asn Asn Thr Val Thr Ala Gln Ile Gly  
130 135 140

Ile Val Ala Val Val Arg Gly Ser Leu Phe Phe Phe Pro Leu Pro Leu  
145 150 155 160

Leu Ile Lys Arg Leu Ala Phe Cys His Ser Asn Val Leu Ser His Ser  
165 170 175



Met Phe Lys Ile Ser Cys Asp Lys Asp Leu Gln Ala Val Gly Gly Lys  
305 310 315 320

<213> Homo Sapien

actatggtcc agaggctgtg

20

<213> Homo Sapien

atcacctatg tgccgcctct

20

<213> Homo sapiens

ggcacgagaa ttaaaaccct cagcaaaaca ggcatagaag ggacatacct taaagtaata 60



```

aaaaccacct atgacaagcc cacagccaac ataatactaa atgggggaaaa gttagaagca 120
tttcctctga gaactgcaac aataaatata aggatgctgg attttgtcaa atgccttttc 180
tgtgtctgtt gagatgctta tgtgactttg cttttaattc tgtttatgtg attatcacat 240
ttattgactt gctgtgttta gaccggaaga gctgggggtg ttctcaggag ccaccgtgtg 300
ctgcggcagc ttcgggataa cttgaggctg catcactggg gaagaaacac aytccctgtcc 360
gtggcgctga tggctgagga cagagcttca gtgtggcttc tctgcgactg gcttcttcgg 420
ggagtctctc cttcatagtt catccatatg gctccagagg aaaattatat tattttgtta 480
tggatgaaga gtattacgtt gtgcagatat actgcagtgt cttcatctct tgatgtgtga 540
ttgggtaggt tccaccatgt tggcgagat gacatgattt cagtacctgt gtctggctga 600
aaagtgtttg tttgtgaatg gatattgtgg tttctggatc tcatcctctg tgggtggaca 660
gctttctcca ccttgtctga agtgacctgc tgtccagaag tttgatggct gaggagtata 720
ccatcgtgca tgcactcttc atttctctga tttcttctc cctggatgga cagggggagc 780
ggcaagagca acgtgggcac ttctggagac cacaacgact cctctgtgaa gacgcttggg 840
agcaagaggt gcaagtgggt ctgccactgc ttccctgct gcagggggag cggcaagagc 900
aacgtggctg cttggggaga ctacgatgac agcgcttca tggatcccag gtaccacgtc 960
catggagaag atctggacaa gctccacaga gctgcctggt ggggtaaagt ccccagaaag 1020
gatctcatcg tcatgctcag ggacacggat gtgaacaaga gggacaagca aaagaggact 1080
gctctacatc tggcctctgc caatgggaat tcagaagtag taaaactcgt gctggacaga 1140
cgatgtcaac ttaatgtcct tgacaacaaa aagaggacag ctctgacaaa ggccgtacaa 1200
tgccaggaag atgaatgtgc gttaatgttg ctggaacatg gcatgatcc aaatattcca 1260
gatgagtatg gaaataccac tctacactat gctgtctaca atgaagataa attaatggcc 1320
aaagcactgc tcttatacgg tgcctgatgc gaatcaaaaa acaagcatgg cctcacacca 1380
ctgctacttg gtatacatga gcaaaaaacag caagtgggtg aatttttaat caagaaaaaa 1440
gcgaatttaa atgcgctgga tagatatgga agaactgctc tcatacttgc tgtatgttgt 1500
ggatcagcaa gtatagtcag cctctactt gagcaaaatg ttgatgtatc ttctcaagat 1560
ctggaaagac ggccagagag tatgctgttt ctagtcatca tcatgtaatt tgccagttac 1620
tttctgacta caaagaaaaa cagatgttaa aaatctcttc tgaaaacagc aatccagaac 1680
aagacttaaa gctgacatca gaggaagagt cacaaaggct taaaggaagt gaaaacagcc 1740
agccagagct agaagattta tggctattga agaagaatga agaacacgga agtactcatg 1800
tgggattccc agaaaacctg actaacggtg ccgctgctgg caatggtgat ga 1852

```

<210> 531

<211> 879

<212> DNA

<213> Homo sapiens

<400> 531

```

atgcatcttt catttctctgc atttcttct ccttggatgg acagggggag cggcaagagc 60
aacgtgggca cttctggaga ccacaacgac tctctgtga agacgcttgg gagcaagagg 120
tgcaagtggg gctgccactg cttccctctg tgcaggggga gcggaagag caacgtggctc 180
gcttggggag actacgatga cagcgcttc atggatccca ggtaccacgt ccatggagaa 240
gatctggaca agctccacag agctgcctgg tggggtaaag tcccagaaa ggatctcatc 300
gtcatgctca gggacacgga tgtgaacaag agggacaagc aaaagaggac tgctctacat 360
ctggcctctg ccaatgggaa ttcagaagta gtaaaactcg tgctggacag acgatgtcaa 420
cttaatgtcc ttgacaacaa aaagaggaca gctctgacaa aggccgtaca atgccaggaa 480
gatgaatgtg cgtaaatgtt gctggaacat ggcactgatc caaatattcc agatgagtat 540
ggaaatacca ctctacacta tgctgtctac aatgaagata aattaatggc caaagcactg 600
ctcttatacg gtgctgatat cgaatcaaaa aacaagcatg gcctcacacc actgctactt 660
ggtatacatg agcaaaaaa gcaagtgggt aaatttttaa tcaagaaaaa agcgaattta 720
aatgcgctgg atagatatgg aagaactgct ctcatacttg ctgtatgttg tggatcagca 780
agtatagtca gccctctact tgagcaaaaat gttgatgtat cttctcaaga tctggaaaga 840

```



879

<211> 293

&lt;212&gt; PRT

<213> Homo sapiens

<400> 532

Met His Leu Ser Phe Pro Ala Phe Leu Pro Pro Trp Met Asp Arg Gly

5 10 15

Ser Gly Lys Ser Asn Val Gly Thr Ser Gly Asp His Asn Asp Ser Ser  
20 25 30

Val Lys Thr Leu Gly Ser Lys Arg Cys Lys Trp Cys Cys His Cys Phe  
35 40 45

Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Val Val Ala Trp Gly Asp  
50 55 60

Tyr Asp Asp Ser Ala Phe Met Asp Pro Arg Tyr His Val His Gly Glu  
65 70 75 80

Asp Leu Asp Lys Leu His Arg Ala Ala Trp Trp Gly Lys Val Pro Arg  
85 90 95

Lys Asp Leu Ile Val Met Leu Arg Asp Thr Asp Val Asn Lys Arg Asp  
100 105 110

Lys Gln Lys Arg Thr Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser  
115 120 125

Glu Val Val Lys Leu Val Leu Asp Arg Arg Cys Gln Leu Asn Val Leu  
130 135 140

Asp Asn Lys Lys Arg Thr Ala Leu Thr Lys Ala Val Gln Cys Gln Glu  
145 150 155 160

Asp Glu Cys Ala Leu Met Leu Leu Glu His Gly Thr Asp Pro Asn Ile  
165 170 175

Pro Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr Ala Val Tyr Asn Glu  
180 185 190

Asp Lys Leu Met Ala Lys Ala Leu Leu Leu Tyr Gly Ala Asp Ile Glu  
195 200 205

Ser Lys Asn Lys His Gly Leu Thr Pro Leu Leu Leu Gly Ile His Glu  
210 215 220

Gln Lys Gln Gln Val Val Lys Phe Leu Ile Lys Lys Lys Ala Asn Leu







Trp Ala Ser Trp Leu Pro Arg Gly Arg Pro  
260 265

cctccactat tacagcttat aggaaattac aatccacttt acaggcctca aaggttcatt 60  
ctggcgcgagc ggacaggcgt ggcggcgga gccccagcat ccctgcttga ggtccaggag 120



cggagcccg	ggccaactgcc	gcctgatcag	cgcgaccccg	gcccgcgccc	gccccgcccc	180
gcaagatgct	gcccggtgtac	cagggaaggtga	agcccaaccc	gctgcaggac	gcgaacctct	240
gctcacgcgt	gttcttctgg	tggctcaatc	ccttgtttaa	aattggccat	aaacggagat	300
tagaggaaaga	tgatatgtat	tcagtgtctgc	cagaagaocg	ctcacagcac	cttggagagg	360
agttgcaagg	gttctgggat	aaagaagttt	taagagctga	gaatgacgca	cagaagcctt	420
ctttaacaag	agcaatcata	aagtgttact	ggaaatctta	tttagttttg	ggaattttta	480
cgttaattga	ggaaaagtgcc	aaagtaatcc	agcccatatt	tttgggaaaa	attattaatt	540
attttgaaaa	ttatgatccc	atggattctg	tggctttgaa	cacagcgta	gcctatgcca	600
cggtgctgac	tttttgcacg	ctcattttgg	ctatactgca	tcacttatat	ttttatcacg	660
ttcagtgtgc	tgggatgagg	ttacgagtag	ccatgtgcc	tatgatttat	cggaaggcac	720
ttcgtcttag	taacatggcc	atggggaaga	caaccacagg	ccagatagtc	aatctgctgt	780
ccaatgatgt	gaacaagttt	gatcagggtga	cagtgttctt	acacttcctg	tgggcaggac	840
cactgcaggc	gatcgagtg	actgcctac	tctggatgga	gataggaata	tcgtgccttg	900
ctgggatggc	agttctaate	attctcctgc	ccttgcaaa	ctgttttggg	aagttgttct	960
catcactgag	gagtaaaact	gcaactttca	cggatgccag	gatcaggacc	atgaatgaag	1020
ttataactgg	tataaggata	ataaaaatgt	acgcctggga	aaagtcattt	tcaaatctta	1080
ttaccaattt	gagaaagaag	gagatttcca	agattctgag	aagttcctgc	ctcaggggga	1140
tgaatttggc	ttcgtttttc	agtgcgaagc	aaatcatcgt	gtttgtgacc	ttcaccacct	1200
acgtgctcct	cggcagtggt	atcacagcca	gccgcgtgtt	cgtggcagtg	acgtctgatg	1260
gggctgtgcy	gctgacgggt	accctcttct	tccctcagc	cattgagagg	gtgtcagagg	1320
caatcgtcag	catccgaaga	atccagacct	ttttgctact	tgatgagata	tcacagcgca	1380
accgtcagct	gccgtcagat	ggtaaaaaga	tgggtgcattg	gcaggatttt	actgcttttt	1440
gggataaggc	atcagagacc	ccaactctac	aaggcctttc	ctttactgtc	agacctggcg	1500
aattgttagc	tgtggtcggc	cccgtggggag	cagggaagtc	atcactgtta	agtgccgtgc	1560
tcggggaatt	ggccccaagt	cacgggctgg	tcagcgtgca	tggagaatt	gcctatgtgt	1620
ctcagcagcc	ctgggtgttc	tcgggaactc	tgaggagtaa	tattttattt	gggaagaaat	1680
acgaaaagga	acgatatgaa	aaagtcataa	aggcttgtgc	tctgaaaaag	gatttacagc	1740
tgttgaggga	tggatgctg	actgtgatag	gagatcgggg	aaccacgctg	agtggagggc	1800
agaaagcacg	ggtaaacctt	gcaagagcag	tgtatcaaga	tgctgacatc	tatctcctgg	1860
acgatcctct	cagtgcagta	gatgcgggaag	ttagcagaca	cctgttcgaa	ctgtgtattt	1920
gtcaaatttt	gcatgagaag	atcacaattt	tagtgactca	tcagttgcag	tacctcaaag	1980
ctgcaagtca	gattctgata	ttgaaaagtg	gtaaaatggt	gcagaagggg	acttacactg	2040
agttcctaaa	atctgggtata	gattttggct	ccctttttaa	gaaggataat	gaggaaagtg	2100
aacaacctcc	agttccagga	actcccacac	taaggaatcg	taccttctca	gagtcttcgg	2160
tttggctctc	acaatcttct	agacctcct	tgaagatgg	tgctctggag	agccaagata	2220
cagagaatgt	cccagttaca	ctatcagagg	agaaccgttc	tgaaggaaaa	gttgggtttt	2280
aggcctataa	gaattacttc	agagctgggt	ctcactggat	tgtcttcatt	ttccttattc	2340
tcctaaacac	tgacgctcag	gttgccctatg	tgcttcaaga	ttgggtggctt	tcatactggg	2400
caaacaaaca	aagtatgcta	aatgtcactg	taaattggagg	aggaaatgta	accgagaagc	2460
tagatcttaa	ctggtaactta	ggaatttatt	caggtttaac	tgtagctacc	gttctttttg	2520
gcatagcaag	atctctattg	gtattctacg	tccttggttaa	ctcttcacaa	actttgcaca	2580
acaaaatggt	tgagtcaatt	ctgaaagctc	cggattattt	ctttgataga	aatccaatag	2640
gaagaatttt	aaatcgtttc	tccaaagaca	ttggacactt	ggatgatttg	ctgccgctga	2700
cgtttttaga	tttcatccag	acatttgctac	aagtggttgg	tgtggtctct	gtggctgtgg	2760
cgtgatttc	ttggatcgca	atacccttgg	ttcccttgg	aatcattttc	atttttcttc	2820
ggcgatattt	tttggaaaacg	tcaagagatg	tgaagcgct	ggaatctaca	actcggagtc	2880
cagtgttttc	ccacttgtca	tcttctctcc	aggggctctg	gacctccgg	gcatacaaa	2940
cagaagagag	gtgtcaggaa	ctgtttgatg	cacaccagga	tttacattca	gaggcttggg	3000
tcttgttttt	gacaacgtcc	cgctgggttcg	cgtccgtct	ggatgccatc	tgtgccatgt	3060
ttgtcatcat	cgttgccctt	gggtccctga	ttctggcaaa	aactctggat	gccgggcagg	3120
ttggtttggc	actgtcctat	qccctcacgc	tcatqqqqat	qtttcagtqq	tgtqttcqac	3180



aaagtgctga agttgagaat atgatgatct cagtagaaaagggtcattgaa tacacagacc 3240  
 ttgaaaaaga agcaccttgg gaatatcaga aacgcccacc accagcctgg ccccatgaag 3300  
 gagtgataat ctttgacaat gtgaacttca tgtacagtc aggtgggcct ctggtactga 3360  
 agcatctgac agcactcatt aaatcacaag aaaagggttg cattgtggga agaaccggag 3420  
 ctggaaaaag ttccctcatc tcagcccttt ttagattgtc agaaccgaa ggtaaaattt 3480  
 ggattgataa gatcttgaca actgaaattg gacttcacga ttaaggaag aaaatgtcaa 3540  
 tcatacctca ggaacctgtt ttgttactg gaacaatgag gaaaaacctg gatcccttta 3600  
 atgagcacac ggatgaggaa ctgtggaatg cettacaaga ggtacaactt aaagaaacca 3660  
 ttgaagatct tcctggtaaa atggatactg aattagcaga atcaggatcc aatttttagtg 3720  
 ttggacaaag acaactgggtg tgccttgcca gggcaattct caggaaaaat cagatattga 3780  
 ttattgatga agcgacggca aatgtggatc caagaactga tgagttaata caaaaaaat 3840  
 ccgggagaaa tttgccact gcaccgtgt aaccattgca cacagattga acaccattat 3900  
 tgacagcgac aagataatgg ttttagattc aggaagactg aaagaatatg atgagcgtga 3960  
 tgttttctg caaaataaag agagcctatt ttacaagatg gtgcaacaac tgggcaaggc 4020  
 agaagcgcgt gccctcactg aaacagcaaa acagggtatac ttcaaaagaa attatccaca 4080  
 tattggctac actgaccaca tggttacaaa cacttccaat ggacagccct cgaccttaac 4140  
 tattttcgag acagcactgt gaatccaacc aaaatgtcaa gtccgttccg aaggcatttg 4200  
 ccactagttt ttggactatg taaaccacat tgtacttttt tttacttttg caacaaatat 4260  
 ttatacatat aagatgctag ttcatattgaa tatttctccc aacttatcca aggatctcca 4320  
 gctctaacaa aatggtttat tttattttaa atgtcaatag ttgtttttta aaatccaaat 4380  
 cagaggtgca ggccaccagt taaatgccgt ctatcagggt ttgtgcctta agagactaca 4440  
 gagtcaaagc tcatttttaa aggagtagga cagagttgtc acaggttttt gttgttgttt 4500  
 ttattgcccc caaaattaca tgttaatttc catttatatc agggattcta tttacttgaa 4560  
 gactgtgaag ttgccatttt gtctcattgt tttctttgac ataactagga tccattattt 4620  
 cccctgaagg cttcttggtta gaaaatagta cagttacaac caataggaac acaaaaaaga 4680  
 aaaagtttgt gacattgtag tagggagtgt gtacccctta ctcccatca aaaaaaaaaa 4740  
 tggatacatg gttaaaggat agaagggcaa tattttatca tatgttctaa aagagaagga 4800  
 agagaaaata ctactttctc aaaatggaag cccttaaagg tgctttgata ctgaaggaca 4860  
 caaatgtgac cgtccatcct cctttagagt tgcatgactt ggacacggta actgttgacg 4920  
 ttttagactc agcattgtga cacttcccaa gaaggccaaa cctctaaccg acattcctga 4980  
 aatacgtggc attattcttt tttggatttc tcatttatgg aaggctaacc ctctgttgac 5040  
 tgtaagcctt ttggtttggg ctgtattgaa atcctttcta aattgcatga ataggctctg 5100  
 ctaacgtgat gagacaaact gaaaattatt gcaagcattg actataatta tgcagtacgt 5160  
 tctcaggatg catccagggg ttcatatttc tgagcctgtc caggttagtt tactcctgac 5220  
 cactaatagc attgtcattt gggtctttctg ttgaatgaat caacaaacca caatacttcc 5280  
 tgggaccttt tgtactttat ttgaactatg agtctttaat ttttctgat gatggtggct 5340  
 gtaatatggt gagttcagtt tactaaagggt tttactatta tgggttgaag tggagtctca 5400  
 tgacctctca gaataagggtg tcacctccct gaaattgcat atatgtatat agacatgcac 5460  
 acgtgtgcat ttgtttgtat acatatattt gtccctcgta tagcaagttt tttgctcatc 5520  
 agcagagagc aacagatggt ttattgagtg aagccttaaa aagcacacac cacacacagc 5580  
 taactgccaa aatacattga ccgtagtagc tgttcaactc ctagtactta gaaatacacg 5640  
 tatggttaat gttcagtcca acaaaccaca cacagtaaat gtttattaat agtcatgggt 5700  
 cgtatttttag gtgactgaaa ttgcaacagt gatcataatg aggtttgtta aaatgatagc 5760  
 tatattcaaa atgtctatat gtttatttgg acttttgagg ttaaagacag tcatataaac 5820  
 gtctgttttc tgttttaatg ttatcataga attttttaat gaaactaaat tcaattgaaa 5880  
 taaatgatag ttttcatctc caaaaaaaaaa aaaaaaaagg gcggccgctc gagtctagag 5940  
 ggcccgttta aaccgcgtga tcagcctcga ctgtgccttc tagttgccag ccactctgtt 6000  
 tttgccctc ccccggtcct tccttgacct tggaagggtgc cactccact gtcctttcct 6060  
 aataaaatga ggaaattgca tc 6082



cagtggcgca	gtctcagetc	actgcagcct	ccacctctctg	tgttcaagca	gtcctcctgc	60
ctcagccacc	agactagcag	gtctcccccg	cctctttctt	ggaaggacac	ttgccattgg	120
atttaggacc	cacttggaata	atccaggatg	atgtcttcac	tccaacatcc	tcagtttaaat	180
tccatgtgca	aatacccttt	tcccaaataa	cattcaattc	tttaccagga	aagggtggctc	240
aatcccttgt	ttaaaattgg	ccataaacgg	agattagagg	aagatgatat	gtattcagtg	300
ctgccagaag	accgctcaca	gcaccttgga	gaggagttgc	aagggttctg	ggataaagaa	360
gttttaagag	ctgagaatga	cgcacagaag	ccttcttttaa	caagagcaat	cataaagtgt	420
tactggaaat	cttatttagt	tttgggaatt	tttacgttaa	ttgaggaaag	tgccaaagta	480
atccagccca	tatttttggg	aaaaattatt	aattattttg	aaaattatga	tcccatggat	540
tctgtggctt	tgaacacagc	gtacgcctat	gccacggtgc	tgactttttg	cacgctcatt	600
ttggctatac	tgcatacact	atatttttat	cacgttcagt	gtgctgggat	gagggttacga	660
gtagccatgt	gccatatgat	ttatcggaag	gcacttcgtc	ttagtaacat	ggccatgggg	720
aagacaacca	caggccagat	agtcaatctg	ctgtccaatg	atgtgaacaa	gtttgatcag	780
gtgacagtgt	tcttacactt	cctgtgggca	ggaccactgc	aggcgatcgc	agtgactgcc	840
ctactctgga	tggagatagg	aatatcgtgc	cctgctggga	tggcagttct	aatcattctc	900
ctgcccttgc	aaagctgttt	tgggaagttg	ttctcatcac	tgaggagtaa	aactgcaact	960
ttcacggatg	ccaggatcag	gaccatgaat	gaagttataa	ctggtataag	gataataaaa	1020
atgtacgcct	gggaaaagtc	attttcaa	cttattacca	at ttgagaaa	gaaggagatt	1080
tccaagattc	tgagaagttc	ctgcctcagg	gggatgaatt	tggcttcgtt	tttcagtgca	1140
agcaaaatca	tcgtgtttgt	gaccttcacc	acctacgtgc	tcctcggcag	tgtgatcaca	1200
gccagccgcg	tgttcgtggc	agtgaacgtg	tatggggctg	tgcggctgac	ggttaccctc	1260
ttcttcccct	cagccattga	gaggggtgtc	gaggcaatcg	tcagcatccg	agaatccag	1320
acctttttgc	tacttgatga	gatatacacag	cgcaaccgtc	agctgccgtc	agatggtaaa	1380
aagatgggtg	atgtgcagga	ttttactgct	ttttgggata	aggcatcaga	gaccccaact	1440
ctacaaggcc	tttctcttac	tgtcagacct	ggcgaattgt	tagctgtggt	cggccccgtg	1500
ggagcaggga	agtcatacct	gttaagtgcc	gtgctcgggg	aattggcccc	aagtcacggg	1560
ctggtcagcg	tgcataggaag	aattgcctat	gtgtctcagc	agccctgggt	gttctcggga	1620
actctgagga	gtaataat	at ttgggaag	aaatacga	aggaacgata	tgaaaaagtc	1680
ataaaggctt	gtgctctgaa	aaaggattta	cagctgtttg	aggatggtga	tctgactgtg	1740
ataggagatc	ggggaaccac	gctgagtggg	gggcagaaag	cacgggtaaa	ccttgcaaga	1800
gcagtgtatc	aagatgctga	catctatctc	ctggacgata	ctctcagtgc	agtagatgcg	1860
gaagttagca	gacacttgtt	cgaactgtgt	at ttgtca	ttttgcatga	gaagatcaca	1920
at tttagta	ctcatcagtt	gcagtacctc	aaagctgcaa	gtcagattct	gatattgaaa	1980
gatggtaaaa	tgggtgcagaa	ggggacttac	actgagttcc	taaaatctgg	tatagatttt	2040
ggctcccttt	taaagaagga	taatgaggaa	agtgaacaac	ctccagttcc	aggaactccc	2100
acactaagga	atcgtaacct	ctcagagtct	toggtttggt	ctcaacaatc	ttctagaccc	2160
tccttgaaag	atgggtgctct	ggagagccaa	gatacagaga	atgtcccagt	tacactatca	2220
gaggagaacc	gttctgaagg	aaaagttggt	tttcaggcct	ataagaatta	cttcagagct	2280
ggtgctcact	ggattgtctt	cattttcctt	attctcctaa	acactgcagc	tcaggttgcc	2340
tatgtgcttc	aagattgggtg	gctttcatac	tgggcaaaca	aacaaagtat	gctaaatgtc	2400
actgtaaaatg	gaggagggaaa	tgtaacccgag	aagctagatc	ttaactggta	cttaggaatt	2460



tattcaggtt	taactgtagc	taccgttctt	tttggcatag	caagatctct	attggtattc	2520
tacgtccttg	ttaactcttc	acaaactttg	cacaacaaaa	tgtttgagtc	aattctgaaa	2580
gctccggtat	tattctttga	tagaaatcca	ataggaagaa	ttttaaatcg	tttctccaaa	2640
gacattggac	acttggatga	tttgtctccg	ctgacgtttt	tagatttcat	ccagacattg	2700
ctacaagtgg	ttggtgtggg	ctctgtggct	gtggccgtga	ttccttggat	cgcaataccc	2760
ttggttcccc	ttggaatcat	tttcattttt	cttcggcgat	attttttggg	aacgtcaaga	2820
gatgtgaagc	gcctggaatc	tacaactcgg	agtccagttg	tttcccactt	gtcatcttct	2880
ctccaggggc	tctggaccat	ccgggcatac	aaagcagaag	agaggtgtca	ggaactgttt	2940
gatgcacacc	aggatttaca	ttcagaggct	tggttcttgt	ttttgacaac	gtcccgtcgg	3000
ttcgccgtcc	gtctggatgc	catctgtgcc	atgtttgtca	tcategttgc	ctttgggtcc	3060
ctgattctgg	caaaaaactct	ggatgcgggg	caggttgggt	tggcactgtc	ctatgccctc	3120
acgtctcatg	ggatgtttca	gtggtgtggt	cgacaaagtg	ctgaagtgtg	gaatatgatg	3180
atctcagtag	aaagggtcat	tgaatacaca	gaccttgaaa	aagaagcacc	ttgggaatat	3240
cagaaaagcc	caccaccagc	ctggccccc	gaaggagtga	taatctttga	caatgtgaac	3300
ttcatgtaca	gtccaggtgg	gcctctggta	ctgaagcatc	tgacagcact	cattaaatca	3360
caagaaaaag	ttggcattgt	gggaagaacc	ggagctggaa	aaagttccct	catctcagcc	3420
cttttttagat	tgtcagaacc	cgaaggtaaa	at ttggattg	ataagatctt	gacaactgaa	3480
attggacttc	acgattttaag	gaagaaaatg	tcaatcatac	ctcaggaacc	tgttttgttc	3540
actggaacaa	tgaggaaaaa	cctggatccc	tttaatgagc	acacggatga	ggaactgtgg	3600
aatgccttac	aagagggtaca	acttaaagaa	accattgaag	atcttcctgg	taaaatggat	3660
actgaattag	cagaatcagg	atccaatttt	agtgttggac	aaagacaact	ggtgtgcctt	3720
gccagggcaa	ttctcaggaa	aaatcagata	ttgattattg	atgaagcgac	ggcaaattgt	3780
gatccaagaa	ctgatgagtt	aatacaaaaa	aaaatccggg	agaaatttgc	ccactgcacc	3840
gtgctaacca	ttgcacacag	attgaacacc	attattgaca	gcgacaagat	aatggtttta	3900
gattcaggaa	gactgaaaga	atatgatgag	ccgtatgttt	tgctgcaaaa	taaagagagc	3960
ctatttttaca	agatgggtgca	acaactgggc	aaggcagaag	ccgtgcctt	cactgaaaca	4020
gcaaaacaga	gatgggggtt	caccatgttg	gccaggctgg	tctcaaactc	ctgacctcaa	4080
gtgatccacc	tgccttggcc	tcccaaactg	ctgagattac	aggtgtgagc	caccacgccc	4140
agcctgagta	tacttcaaaa	gaaattatcc	acatattggt	cacactgacc	acatggttac	4200
aaacacttcc	aatggacagc	cctcgacctt	aactattttc	gagacagcac	tgtgaatcca	4260
accaaaatgt	caagtccgtt	ccgaaggcat	ttgccactag	tttttggact	atgtaaacca	4320
cattgtactt	ttttttactt	tggcaacaaa	tattttataca	tacaagatgc	tagtttcattt	4380
gaatattttct	cccaacttat	ccaaggatct	ccagctctaa	caaaatgggt	tattttttatt	4440
taaatgtcaa	tagtkgkttt	ttaaaatcca	aatcagagggt	gcaggccacc	agttaaatgc	4500
cgtctatcag	gttttgtgcc	ttaagagact	acagnagtca	gaagctcatt	tttaaaggag	4560
taggacagag	ttgtcacagg	tttttgttgg	tgtttktatt	gcccccaaaa	ttacatgtta	4620
atttccattt	atatcagggg	attctatttt	cttgaagact	gtgaagtgtc	catttttgtct	4680
cattgttttc	tttgacatam	ctaggatcca	ttattttccc	tgaaggcttc	ttgkagaaaa	4740
tagtacagtt	acaaccaata	ggaactamca	aaaagaaaaa	gtttgtgaca	ttgtagtagg	4800
gagtgtgtac	cccttactcc	ccatcaaaaa	aaaaaatgga	tacatggtta	aaggatagaa	4860
gggcaatatt	ttatcatatg	ttctaaaaga	gaaggaagag	aaaatactac	tttctcaaaa	4920
tggaagccct	taaagggtgct	ttgatactga	aggacacaaa	tgtgaccgtc	catcctcctt	4980
tagagttgca	tgacttggac	acggttaactg	ttgcagtttt	agactcagca	ttgtgacact	5040
tcccaagaag	gccaaacctc	taaccgacat	tcctgaaata	cgtggcatta	ttctttttttg	5100
gattttctcat	ttaggaaggc	taaccctctg	ttgamtgtam	kccttttgggt	ttgggctgta	5160
ttgaaatcct	ttctaaattg	catgaatagg	ctctgctaac	cgtgatgaga	caaactgaaa	5220
attattgcaa	gcattgacta	taattatgca	gtacgttctc	aggatgcata	caggggttca	5280
ttttcatgag	cctgtccagg	ttagtttact	cctgaccact	aatagcattg	tcatttgggc	5340
tttctgttga	atgaatcaac	aaaccacaat	acttcctggg	accttttgtg	ctttattttga	5400
actatgagtc	tttaattttt	cctgatgatg	gtggctgtaa	tatgttgagt	tcagtttact	5460
aaagggtttta	ctattatggt	ttgaaggggag	tctcatgacc	tctcaaaaaa	qqtqcacctc	5520



```
<210> 537
<211> 1229
<212> PRT
<213> Homo sapiens
```

<400> 537																
Met	Leu	Pro	Val	Tyr	Gln	Glu	Val	Lys	Pro	Asn	Pro	Leu	Gln	Asp	Ala	
				5					10					15		
.																
Asn	Leu	Cys	Ser	Arg	Val	Phe	Phe	Trp	Trp	Leu	Asn	Pro	Leu	Phe	Lys	
				20					25					30		
Ile	Gly	His	Lys	Arg	Arg	Leu	Glu	Glu	Asp	Asp	Met	Tyr	Ser	Val	Leu	
				35					40					45		
Pro	Glu	Asp	Arg	Ser	Gln	His	Leu	Gly	Glu	Glu	Leu	Gln	Gly	Phe	Trp	
				50					55					60		
Asp	Lys	Glu	Val	Leu	Arg	Ala	Glu	Asn	Asp	Ala	Gln	Lys	Pro	Ser	Leu	
				65					70					75		
Thr	Arg	Ala	Ile	Ile	Lys	Cys	Tyr	Trp	Lys	Ser	Tyr	Leu	Val	Leu	Gly	
				85					90					95		
Ile	Phe	Thr	Leu	Ile	Glu	Glu	Ser	Ala	Lys	Val	Ile	Gln	Pro	Ile	Phe	
				100					105					110		
Leu	Gly	Lys	Ile	Ile	Asn	Tyr	Phe	Glu	Asn	Tyr	Asp	Pro	Met	Asp	Ser	
				115					120					125		
Val	Ala	Leu	Asn	Thr	Ala	Tyr	Ala	Tyr	Ala	Thr	Val	Leu	Thr	Phe	Cys	
				130					135					140		
Thr	Leu	Ile	Leu	Ala	Ile	Leu	His	His	Leu	Tyr	Phe	Tyr	His	Val	Gln	
				145					150					155		
Cys	Ala	Gly	Met	Arg	Leu	Arg	Val	Ala	Met	Cys	His	Met	Ile	Tyr	Arg	
				165					170					175		







Ala 450	Gly	Lys	Ser	Ser	Leu	Leu 455	Ser	Ala	Val	Leu	Gly 460	Glu	Leu	Ala	Pro
Ser 465	His	Gly	Leu	Val	Ser 470	Val	His	Gly	Arg	Ile 475	Ala	Tyr	Val	Ser	Gln 480
Gln	Pro	Trp	Val	Phe 485	Ser	Gly	Thr	Leu	Arg 490	Ser	Asn	Ile	Leu	Phe 495	Gly
Lys	Lys	Tyr	Glu 500	Lys	Glu	Arg	Tyr	Glu 505	Lys	Val	Ile	Lys	Ala 510	Cys	Ala
Leu	Lys 515	Lys	Asp	Leu	Gln	Leu 520	Leu	Glu	Asp	Gly	Asp 525	Leu	Thr	Val	Ile
Gly 530	Asp	Arg	Gly	Thr	Thr 535	Leu	Ser	Gly	Gly	Gln 540	Lys	Ala	Arg	Val	Asn
Leu 545	Ala	Arg	Ala	Val 550	Tyr	Gln	Asp	Ala	Asp 555	Ile	Tyr	Leu	Leu	Asp 560	Asp
Pro	Leu	Ser	Ala 565	Val	Asp	Ala	Glu	Val 570	Ser	Arg	His	Leu	Phe 575	Glu	Leu
Cys	Ile	Cys 580	Gln	Ile	Leu	His	Glu 585	Lys	Ile	Thr	Ile	Leu 590	Val	Thr	His
Gln 595	Leu	Gln	Tyr	Leu	Lys	Ala 600	Ala	Ser	Gln	Ile	Leu 605	Ile	Leu	Lys	Asp
Gly 610	Lys	Met	Val	Gln	Lys 615	Gly	Thr	Tyr	Thr	Glu	Phe 620	Leu	Lys	Ser	Gly
Ile 625	Asp	Phe	Gly	Ser 630	Leu	Leu	Lys	Lys	Asp 635	Asn	Glu	Glu	Ser	Glu	Gln 640
Pro	Pro	Val	Pro 645	Gly	Thr	Pro	Thr	Leu	Arg 650	Asn	Arg	Thr	Phe 655	Ser	Glu
Ser	Ser	Val 660	Trp	Ser	Gln	Gln	Ser 665	Ser	Arg	Pro	Ser	Leu 670	Lys	Asp	Gly
Ala 675	Leu	Glu	Ser	Gln	Asp	Thr 680	Glu	Asn	Val	Pro	Val	Thr 685	Leu	Ser	Glu
Glu 690	Asn	Arg	Ser	Glu	Gly 695	Lys	Val	Gly	Phe	Gln 700	Ala	Tyr	Lys	Asn	Tyr
Phe 705	Arg	Ala	Gly	Ala 710	His	Trp	Ile	Val	Phe 715	Ile	Phe	Leu	Ile	Leu	Leu 720



Asn	Thr	Ala	Ala	Gln	Val	Ala	Tyr	Val	Leu	Gln	Asp	Trp	Trp	Leu	Ser				
725								730				735							
Tyr	Trp	Ala	Asn	Lys	Gln	Ser	Met	Leu	Asn	Val	Thr	Val	Asn	Gly	Gly				
740								745				750							
Gly	Asn	Val	Thr	Glu	Lys	Leu	Asp	Leu	Asn	Trp	Tyr	Leu	Gly	Ile	Tyr				
755								760				765							
Ser	Gly	Leu	Thr	Val	Ala	Thr	Val	Leu	Phe	Gly	Ile	Ala	Arg	Ser	Leu				
770								775				780							
Leu	Val	Phe	Tyr	Val	Leu	Val	Asn	Ser	Ser	Gln	Thr	Leu	His	Asn	Lys				
785				790								795				800			
Met	Phe	Glu	Ser	Ile	Leu	Lys	Ala	Pro	Val	Leu	Phe	Phe	Asp	Arg	Asn				
				805								810				815			
Pro	Ile	Gly	Arg	Ile	Leu	Asn	Arg	Phe	Ser	Lys	Asp	Ile	Gly	His	Leu				
				820								825				830			
Asp	Asp	Leu	Leu	Pro	Leu	Thr	Phe	Leu	Asp	Phe	Ile	Gln	Thr	Leu	Leu				
				835								840				845			
Gln	Val	Val	Gly	Val	Val	Ser	Val	Ala	Val	Ala	Val	Ile	Pro	Trp	Ile				
				850								855				860			
Ala	Ile	Pro	Leu	Val	Pro	Leu	Gly	Ile	Ile	Phe	Ile	Phe	Leu	Arg	Arg				
865								870								875		880	
Tyr	Phe	Leu	Glu	Thr	Ser	Arg	Asp	Val	Lys	Arg	Leu	Glu	Ser	Thr	Thr				
				885								890				895			
Arg	Ser	Pro	Val	Phe	Ser	His	Leu	Ser	Ser	Ser	Leu	Gln	Gly	Leu	Trp				
				900								905				910			
Thr	Ile	Arg	Ala	Tyr	Lys	Ala	Glu	Glu	Arg	Cys	Gln	Glu	Leu	Phe	Asp				
				915								920				925			
Ala	His	Gln	Asp	Leu	His	Ser	Glu	Ala	Trp	Phe	Leu	Phe	Leu	Thr	Thr				
				930								935				940			
Ser	Arg	Trp	Phe	Ala	Val	Arg	Leu	Asp	Ala	Ile	Cys	Ala	Met	Phe	Val				
945								950								955		960	
Ile	Ile	Val	Ala	Phe	Gly	Ser	Leu	Ile	Leu	Ala	Lys	Thr	Leu	Asp	Ala				
				965								970				975			
Gly	Gln	Val	Gly	Leu	Ala	Leu	Ser	Tyr	Ala	Leu	Thr	Leu	Met	Gly	Met				
				980								985				990			



```

Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val Glu Asn Met Met Ile
    995                      1000                      1005

Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala Pro
    1010                      1015                      1020

Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu Gly Val
    1025                      1030                      1035                      1040

Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly Gly Pro Leu
    1045                      1050                      1055

Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu Lys Val Gly
    1060                      1065                      1070

Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile Ser Ala Leu
    1075                      1080                      1085

Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp Lys Ile Leu
    1090                      1095                      1100

Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met Ser Ile Ile
    1105                      1110                      1115                      1120

Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp
    1125                      1130                      1135

Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu
    1140                      1145                      1150

Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr
    1155                      1160                      1165

Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu
    1170                      1175                      1180

Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile
    1185                      1190                      1195                      1200

Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln
    1205                      1210                      1215

Lys Lys Ser Gly Arg Asn Leu Pro Thr Ala Pro Cys
    1220                      1225

```

&lt;210&gt; 538

&lt;211&gt; 1262

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;400&gt; 538



Met	Tyr	Ser	Val	Leu	Pro	Glu	Asp	Arg	Ser	Gln	His	Leu	Gly	Glu	Glu	
				5					10					15		
Leu	Gln	Gly	Phe	Trp	Asp	Lys	Glu	Val	Leu	Arg	Ala	Glu	Asn	Asp	Ala	
				20					25					30		
Gln	Lys	Pro	Ser	Leu	Thr	Arg	Ala	Ile	Ile	Lys	Cys	Tyr	Trp	Lys	Ser	
				35					40					45		
Tyr	Leu	Val	Leu	Gly	Ile	Phe	Thr	Leu	Ile	Glu	Glu	Ser	Ala	Lys	Val	
				50					55					60		
Ile	Gln	Pro	Ile	Phe	Leu	Gly	Lys	Ile	Ile	Asn	Tyr	Phe	Glu	Asn	Tyr	
65					70					75					80	
Asp	Pro	Met	Asp	Ser	Val	Ala	Leu	Asn	Thr	Ala	Tyr	Ala	Tyr	Ala	Thr	
				85					90					95		
Val	Leu	Thr	Phe	Cys	Thr	Leu	Ile	Leu	Ala	Ile	Leu	His	His	Leu	Tyr	
				100					105					110		
Phe	Tyr	His	Val	Gln	Cys	Ala	Gly	Met	Arg	Leu	Arg	Val	Ala	Met	Cys	
				115					120					125		
His	Met	Ile	Tyr	Arg	Lys	Ala	Leu	Arg	Leu	Ser	Asn	Met	Ala	Met	Gly	
				130					135					140		
Lys	Thr	Thr	Thr	Gly	Gln	Ile	Val	Asn	Leu	Leu	Ser	Asn	Asp	Val	Asn	
145					150					155					160	
Lys	Phe	Asp	Gln	Val	Thr	Val	Phe	Leu	His	Phe	Leu	Trp	Ala	Gly	Pro	
				165					170					175		
Leu	Gln	Ala	Ile	Ala	Val	Thr	Ala	Leu	Leu	Trp	Met	Glu	Ile	Gly	Ile	
				180					185					190		
Ser	Cys	Leu	Ala	Gly	Met	Ala	Val	Leu	Ile	Ile	Leu	Leu	Pro	Leu	Gln	
				195					200					205		
Ser	Cys	Phe	Gly	Lys	Leu	Phe	Ser	Ser	Leu	Arg	Ser	Lys	Thr	Ala	Thr	
				210					215					220		
Phe	Thr	Asp	Ala	Arg	Ile	Arg	Thr	Met	Asn	Glu	Val	Ile	Thr	Gly	Ile	
225					230					235					240	
Arg	Ile	Ile	Lys	Met	Tyr	Ala	Trp	Glu	Lys	Ser	Phe	Ser	Asn	Leu	Ile	
				245					250					255		
Thr	Asn	Leu	Arg	Lys	Lys	Glu	Ile	Ser	Lys	Ile	Leu	Arg	Ser	Ser	Cys	
				260					265					270		







Ile	Leu	Val	Thr	His	Gln	Leu	Gln	Tyr	Leu	Lys	Ala	Ala	Ser	Gln	Ile
545					550					555					560
Leu	Ile	Leu	Lys	Asp	Gly	Lys	Met	Val	Gln	Lys	Gly	Thr	Tyr	Thr	Glu
				565					570					575	
Phe	Leu	Lys	Ser	Gly	Ile	Asp	Phe	Gly	Ser	Leu	Leu	Lys	Lys	Asp	Asn
			580					585					590		
Glu	Glu	Ser	Glu	Gln	Pro	Pro	Val	Pro	Gly	Thr	Pro	Thr	Leu	Arg	Asn
		595					600					605			
Arg	Thr	Phe	Ser	Glu	Ser	Ser	Val	Trp	Ser	Gln	Gln	Ser	Ser	Arg	Pro
	610					615					620				
Ser	Leu	Lys	Asp	Gly	Ala	Leu	Glu	Ser	Gln	Asp	Thr	Glu	Asn	Val	Pro
625					630					635					640
Val	Thr	Leu	Ser	Glu	Glu	Asn	Arg	Ser	Glu	Gly	Lys	Val	Gly	Phe	Gln
				645					650					655	
Ala	Tyr	Lys	Asn	Tyr	Phe	Arg	Ala	Gly	Ala	His	Trp	Ile	Val	Phe	Ile
			660					665					670		
Phe	Leu	Ile	Leu	Leu	Asn	Thr	Ala	Ala	Gln	Val	Ala	Tyr	Val	Leu	Gln
		675					680					685			
Asp	Trp	Trp	Leu	Ser	Tyr	Trp	Ala	Asn	Lys	Gln	Ser	Met	Leu	Asn	Val
	690					695					700				
Thr	Val	Asn	Gly	Gly	Gly	Asn	Val	Thr	Glu	Lys	Leu	Asp	Leu	Asn	Trp
705					710					715					720
Tyr	Leu	Gly	Ile	Tyr	Ser	Gly	Leu	Thr	Val	Ala	Thr	Val	Leu	Phe	Gly
			725						730					735	
Ile	Ala	Arg	Ser	Leu	Leu	Val	Phe	Tyr	Val	Leu	Val	Asn	Ser	Ser	Gln
			740					745					750		
Thr	Leu	His	Asn	Lys	Met	Phe	Glu	Ser	Ile	Leu	Lys	Ala	Pro	Val	Leu
		755					760					765			
Phe	Phe	Asp	Arg	Asn	Pro	Ile	Gly	Arg	Ile	Leu	Asn	Arg	Phe	Ser	Lys
		770				775					780				
Asp	Ile	Gly	His	Leu	Asp	Asp	Leu	Leu	Pro	Leu	Thr	Phe	Leu	Asp	Phe
785					790					795					800
Ile	Gln	Thr	Leu	Leu	Gln	Val	Val	Gly	Val	Val	Ser	Val	Ala	Val	Ala
				805					810					815	



Val	Ile	Pro	Trp	Ile	Ala	Ile	Pro	Leu	Val	Pro	Leu	Gly	Ile	Ile	Phe	
			820						825						830	
Ile	Phe	Leu	Arg	Arg	Tyr	Phe	Leu	Glu	Thr	Ser	Arg	Asp	Val	Lys	Arg	
		835					840					845				
Leu	Glu	Ser	Thr	Thr	Arg	Ser	Pro	Val	Phe	Ser	His	Leu	Ser	Ser	Ser	
	850					855					860					
Leu	Gln	Gly	Leu	Trp	Thr	Ile	Arg	Ala	Tyr	Lys	Ala	Glu	Glu	Arg	Cys	
865					870					875					880	
Gln	Glu	Leu	Phe	Asp	Ala	His	Gln	Asp	Leu	His	Ser	Glu	Ala	Trp	Phe	
			885						890						895	
Leu	Phe	Leu	Thr	Thr	Ser	Arg	Trp	Phe	Ala	Val	Arg	Leu	Asp	Ala	Ile	
			900					905					910			
Cys	Ala	Met	Phe	Val	Ile	Ile	Val	Ala	Phe	Gly	Ser	Leu	Ile	Leu	Ala	
	915						920					925				
Lys	Thr	Leu	Asp	Ala	Gly	Gln	Val	Gly	Leu	Ala	Leu	Ser	Tyr	Ala	Leu	
	930					935					940					
Thr	Leu	Met	Gly	Met	Phe	Gln	Trp	Cys	Val	Arg	Gln	Ser	Ala	Glu	Val	
945					950					955					960	
Glu	Asn	Met	Met	Ile	Ser	Val	Glu	Arg	Val	Ile	Glu	Tyr	Thr	Asp	Leu	
				965					970					975		
Glu	Lys	Glu	Ala	Pro	Trp	Glu	Tyr	Gln	Lys	Arg	Pro	Pro	Pro	Ala	Trp	
		980						985						990		
Pro	His	Glu	Gly	Val	Ile	Ile	Phe	Asp	Asn	Val	Asn	Phe	Met	Tyr	Ser	
	995						1000						1005			
Pro	Gly	Gly	Pro	Leu	Val	Leu	Lys	His	Leu	Thr	Ala	Leu	Ile	Lys	Ser	
	1010					1015						1020				
Gln	Glu	Lys	Val	Gly	Ile	Val	Gly	Arg	Thr	Gly	Ala	Gly	Lys	Ser	Ser	
1025					1030						1035					1040
Leu	Ile	Ser	Ala	Leu	Phe	Arg	Leu	Ser	Glu	Pro	Glu	Gly	Lys	Ile	Trp	
				1045					1050						1055	
Ile	Asp	Lys	Ile	Leu	Thr	Thr	Glu	Ile	Gly	Leu	His	Asp	Leu	Arg	Lys	
		1060						1065						1070		
Lys	Met	Ser	Ile	Ile	Pro	Gln	Glu	Pro	Val	Leu	Phe	Thr	Gly	Thr	Met	
	1075						1080						1085			



```

Arg Lys Asn Leu Asp Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp
1090                      1095                      1100

Asn Ala Leu Gln Glu Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro
1105                      1110                      1115                      1120

Gly Lys Met Asp Thr Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val
                      1125                      1130                      1135

Gly Gln Arg Gln Leu Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn
                      1140                      1145                      1150

Gln Ile Leu Ile Ile Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr
                      1155                      1160                      1165

Asp Glu Leu Ile Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr
1170                      1175                      1180

Val Leu Thr Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys
1185                      1190                      1195                      1200

Ile Met Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr
                      1205                      1210                      1215

Val Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln
                      1220                      1225                      1230

Leu Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Arg
1235                      1240                      1245

Trp Gly Phe Thr Met Leu Ala Arg Leu Val Ser Asn Ser
1250                      1255                      1260

<210> 539
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> Made in a lab

<400> 539
Cys Leu Ser His Ser Val Ala Val Val Thr
 1             5             10

```

```

<210> 540
<211> 9
<212> PRT
<213> Artificial Sequence

```

```

<220>

```



<400> 540

<210> 541

<211> 14

<212> PRT

<213> Homo sapiens

<400> 541

<210> 542

<211> 15

<212> PRT

<213> Homo sapiens

<400> 542

<210> 543

<211> 12

<212> PRT

<213> Homo sapiens

<400> 543

<210> 544

<211> 18

<212> PRT

<213> Homo sapiens

<400> 544

Met Thr

<210> 545

<211> 18



<212> PRT

<213> Homo sapiens

<400> 545

Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala  
                   5                  10                  15

Ser Val

<210> 546

<211> 29

<212> PRT

<213> Homo sapiens

<400> 546

Phe Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly  
                   5                  10                  15

Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg Met  
                   20                  25

<210> 547

<211> 58

<212> PRT

<213> Homo sapiens

<400> 547

Val Ala Glu Glu Ala Ala Leu Gly Pro Thr Glu Pro Ala Glu Gly Leu  
                   5                  10                  15

Ser Ala Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg Ala Arg Leu  
                   20                  25                  30

Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His Gln Leu Cys  
                   35                  40                  45

Cys Arg Met Pro Arg Thr Leu Arg Arg Leu  
                   50                  55

<210> 548

<211> 18

<212> PRT

<213> Homo sapiens

<400> 548

Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu  
                   5                  10                  15



Glu Cys

<210> 549

<211> 18

<212> PRT

<213> Homo sapiens

<400> 549

Leu Glu Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro Asp His Cys Arg  
                           5                          10                          15

Gln Ala

<210> 550

<211> 14

<212> PRT

<213> Homo sapiens

<400> 550

Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe  
                           5                          10

<210> 551

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 551

Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala

<210> 552

<211> 2577

<212> DNA

<213> Homo sapiens

<400> 552

agcatatgta acatgacctg tgcttcagtg ttcttttgtg atcaaaaatt ccttactttt 60  
 agttttttat ctatggtaga accacccaga gcaggggtcc tcaactccca ggccacagac 120  
 tcataaccagt ccacggacta ttatgaacca caccacacag gaggaggtga gcactaggca 180  
 agccaaggaa gcttcacctg tacttacagc cacacgccat ggctcatatt acagcctgaa 240  
 ctctgcctcc actcagatca gtgataacat tagaaactca ttggagcacg aaccctgttg 300



```

tgaactgcct atccgaagga tctaggttgt gtgcttcgta tgagaatcta atgccagatg 360
atctatcatt gtctcacttt gccccagat aagaccatct agttgcagaa aaataagctc 420
agagcttcca ctgattctac attatggata tgtgccgccg aagcaagcac aaagccctac 480
ttttacacat gcttagtgat gcttcatgga caaggcttgg ctctgttgag tccaactaac 540
ctacctgaga ttctgagatt tctcttcaat ggcttcctgt gagctagagt ttgaaaatat 600
cttaaaatct tgagctagag atggaagtag cttggacgat tttcattatc atgtaaatcg 660
ggtcactcaa ggggccaaacc acagctggga gccactgctc aggggaaggt tcatatggga 720
ctttctactg cccaaggttc tatacaggat ataaagggtgc ctcacagtat agatctggta 780
gcaaagaaga agaaacaaac actgatctct ttctgccacc cctctgacct tttggaactc 840
ctctgacctt ttagaacaag cctacctaat atctgctaga gaaaagacca acaacggcct 900
caaaggatct cttaccatga aggtctcagc taattcttgg ctaagatgtg ggttccacat 960
taggttctga atatgggggg aagggtcaat ttgctcattt tgtgtgtgga taaagtcagg 1020
atgcccaggg gccagagcag ggggctgctg ctttggaac aatggctgag catataacca 1080
taggtatggg aacaaaaaac atcaaagtca ctgtatcaat tgccatgaag actcgaggga 1140
cctgaatcta ccgattcatc ttaaggcagc aggaccagtt tgagtggcaa caatgcagca 1200
gcagaatcaa tggaaacaac agaatgattg caatgtcctt ttttttctcc tccttctgac 1260
ttgataaaag ggaccgtctt ccttggattt agtgaacccc tttggttcct gaaaaattca 1320
aggagtatct aggacatagt cccagaaga cagtacaaga ctttctgata aactggacat 1380
ttcaagcccc aaataactaa tcagaaaaat caaagatgtg atactatttt ttatcccatg 1440
catagggtgt acacttggat caaatgaaca atgttgggat ctytatggat aaaggtctta 1500
aaagtcctga gataaagaat cctgcaccca ctggtacttc taacttgtct tgttttttgt 1560
ctgwtttctg gctgatgcag gggactaact cactgccacg cgaaaactac ctgaactgaa 1620
ctatgacatc tcacctgata tgtaagatgt aactgttata attattttaa acctcaattt 1680
agcattaact agccttttaa tgtaaacact tacacattat gaygactaga aacagcatac 1740
tctctggccg tctgtccaga tagatcttga gaagatacat caatgttttg ctcaagtaga 1800
aggctgacta tacttgccga tccacaacat acagcaagta tgagagcagt tctaaaatga 1860
cagagatagg aacagtaata aagttattkt aaaagcta attgatatact ttaccaattt 1920
aacatcttgc ctgtccgtgc agaatcaaac atttacatgc actaaaagac ataagcatct 1980
tcagtgtcga agtgttcate tttgtaaaat accaccaagg ttaaaaggaa gggacaaaaa 2040
aaaaaaaccc tcttatctca gtgggggtatt gcatagcaga agctactaat ttgaagtcct 2100
ttgatggaca agaaacaata ttaggggccac ttatctgaaa tgaacaaaga tttgaagtga 2160
gatttcatca cagcttccct agactgatat gctgtaatag aaaatcagct aggggggtaaa 2220
ataaataaga gctctctgca tgctgaaagc aagtaagatt aataataatg gtaagaatag 2280
tagtcacagg agtttcagtt aatgatgcca ataagcatgt gctaggcact gaattaaatg 2340
ccacatatat ctttcttatg cgcagcaaac tttgaaggat atattctcct acttttcata 2400
tatgacaaca tatttggttg taaataacgt tccaagggtc acacacctag caagtaagaa 2460
agttaggaat taaaccagat attgtgtgaa tctaaagcct aacttttttc tctttatcac 2520
ccacctacgg cttgtcttca ttaaaggaaa agtgtatcca cttaaaaaaa aaaaaaa 2577

```

<210> 553

<211> 58

<212> PRT

<213> Homo sapiens

<400> 553

Ser Ile Cys Asn Met Thr Cys Ala Ser Val Phe Phe Cys Asp Gln Lys  
5 10 15

Phe Leu Thr Phe Ser Phe Leu Ser Met Val Glu Pro Pro Arg Ala Gly  
20 25 30



```
<210> 556
<211> 81
<212> PRT
```



<400> 556

Ser Pro Arg Thr Ile Met Asn His Thr Thr Gln Glu Glu Val Ser Thr  
20 25 30

Arg Gln Ala Lys Glu Ala Ser Pro Val Leu Thr Ala Thr Arg His Gly  
35 40 45

Ser Tyr Tyr Ser Leu Asn Ser Ala Ser Thr Gln Ile Ser Asp Asn Ile  
50 55 60

Arg Asn Ser Leu Glu His Glu Pro Cys Cys Glu Leu Pro Ile Arg Arg  
65 70 75 80

Ile

<210> 557

<211> 54

<212> PRT

<213> Homo sapiens

<400> 557

Ser Leu Ser Ala Thr Pro Leu Thr Leu Trp Asn Ser Ser Asp Pro Leu  
5 10 15

Glu Gln Ala Tyr Leu Ile Ser Ala Arg Glu Lys Thr Asn Asn Gly Leu  
20 25 30

Lys Gly Ser Leu Thr Met Lys Val Ser Ala Asn Ser Trp Leu Arg Cys  
35 40 45

Gly Phe His Ile Arg Phe  
50

<210> 558

<211> 77

&lt;212&gt; PRT

<213> Homo sapiens

<220>

&lt;221&gt; VARIANT

<222> (1) ... (77)

<223> Xaa = Any amino acid



Asn Asp Arg Asp Arg Asn Ser Asn Lys Val Ile Xaa Lys Ala Asn Leu  
5 10 15

Ile Tyr Phe Thr Asn Leu Thr Ser Cys Leu Ser Val Gln Asn Gln Thr  
20 25 30

Phe Thr Cys Thr Lys Arg His Lys His Leu Gln Cys Ser Ser Val His  
35 40 45

Leu Cys Lys Ile Pro Pro Arg Leu Lys Gly Arg Asp Lys Lys Lys Lys  
50 55 60

Pro Ser Tyr Leu Ser Gly Val Leu His Ser Arg Ser Tyr  
65 70 75

<211> 50

<212> PRT

<213> Homo sapiens

Thr Leu Pro Pro Leu Arg Ser Val Ile Thr Leu Glu Thr His Trp Ser  
5 10 15

Thr Asn Pro Val Val Asn Cys Leu Ser Glu Gly Ser Arg Leu Cys Ala  
20 25 30

Ser Tyr Glu Asn Leu Met Pro Asp Asp Leu Ser Leu Ser His Phe Ala  
35 40 45

Pro Arg  
50

<211> 56

<212> PRT

<213> Homo sapiens

Ile Gly Ser Leu Lys Gly Pro Thr Thr Ala Gly Ser His Cys Ser Gly  
5 10 15

Glu Gly Ser Tyr Gly Thr Phe Tyr Cys Pro Arg Phe Tyr Thr Gly Tyr  
20 25 30

Lys Gly Ala Ser Gln Tyr Arg Ser Gly Ser Lys Glu Glu Glu Thr Asn  
35 40 45



Leu Ser Ser Gly Asp Tyr Val Leu Asp Thr Pro  
50 55



<211> 79

&lt;212&gt; PRT

<213> Homo sapiens

Cys Phe Leu Phe Pro Tyr Leu Trp Leu Tyr Ala Gln Pro Leu Phe Pro  
5 10 15

Lys Gln Gln Pro Pro Ala Leu Ala Pro Gly His Pro Asp Phe Ile His  
20 25 30

Thr Gln Asn Glu Gln Ile Asp Pro Ser Pro His Ile Gln Asn Leu Met  
35 40 45

Trp Asn Pro His Leu Ser Gln Glu Leu Ala Glu Thr Phe Met Val Arg  
50 55 60

Asp Pro Leu Arg Pro Leu Leu Val Phe Ser Leu Ala Asp Ile Arg  
65 70 75

<211> 64

<212> PRT

<213> Homo sapiens

Ala Cys Ser Lys Gly Ser Glu Glu Phe Gln Arg Val Arg Gly Val Ala  
5 10 15

Glu Arg Asp Gln Cys Leu Phe Leu Leu Leu Cys Tyr Gln Ile Tyr Thr  
20 25 30

Val Arg His Leu Tyr Ile Leu Tyr Arg Thr Leu Gly Ser Arg Lys Ser  
35 40 45

His Met Asn Leu Pro Leu Ser Ser Gly Ser Gln Leu Trp Leu Ala Pro  
50 55 60

<211> 57

<212> PRT

<213> Homo sapiens

<220>

&lt;221&gt; VARIANT

$\langle 222 \rangle$  (1) ... (57)

<223> Xaa = Any amino acid



Leu	Tyr	Tyr	Cys	Ser	Tyr	Leu	Cys	His	Phe	Arg	Thr	Ala	Leu	Ile	Leu	
				5					10					15		
Ala	Val	Cys	Cys	Gly	Ser	Ala	Ser	Ile	Val	Ser	Leu	Leu	Leu	Glu	Gln	
				20					25					30		
Asn	Ile	Asp	Val	Ser	Ser	Gln	Asp	Leu	Ser	Gly	Gln	Thr	Ala	Arg	Glu	
				35					40					45		
Tyr	Ala	Val	Ser	Ser	Xaa	His	Asn	Val								
				50					55							

<213> Homo sapiens

Ile	Leu	Leu	Glu	Phe	Phe	Arg	Asn	Gln	Arg	Gly	Ser	Leu	Asn	Pro	Arg
				5					10					15	
Lys	Thr	Val	Pro	Phe	Ile	Lys	Ser	Glu	Gly	Gly	Glu	Lys	Lys	Gly	His
			20					25					30		
Cys	Asn	His	Ser	Val	Val	Ser	Ile	Asp	Ser	Ala	Ala	Ala	Leu	Leu	Pro
		35					40					45			
Leu	Lys	Leu	Val	Leu	Leu	Pro									
	50					55									

<213> Homo sapiens

[illegible]



```

<400> 568
Lys Val Gly Glu Tyr Ile Leu Gln Ser Leu Leu Arg Ile Arg Lys Ile
      5                      10                      15
Tyr Val Ala Phe Asn Ser Val Pro Ser Thr Cys Leu Leu Ala Ser Leu
      20                      25                      30
Thr Glu Thr Pro Val Thr Thr Ile Leu Thr Ile Ile Ile Asn Leu Thr
      35                      40                      45
Cys Phe Gln His Ala Glu Ser Ser Tyr Leu Phe Tyr Pro Leu Ala Asp
      50                      55                      60
Phe Leu Leu Gln His Ile Ser Leu Gly Lys Leu
      65                      70                      75

```

<400>	569						
gcatccagag	tgggtggactg	gttacaggct	atgaacctac	actgatgcgg	caccaccacc	60	
cagagtccac	rggttatgtt	ggttcacatt	tactottgct	gtggtatggt	ctataggttt	120	
ggacagatgt	ccgataatcc	tttttacatt	ttggcatcct	tgggtagctc	gtcttgtagg	180	
aatggacttg	cttcaaagtg	gaggcaggca	gacccctcag	acgggtatat	ggagccctgt	240	
tttcagttgc	ttttctaatt	ctctcttata	gtttacctca	aaatcttcct	gagggtctcgc	300	
ttccttttaa	aatccttgtc	tactttgcag	catcactctg	acactcccat	tgattcctca	360	
gcacctactg	actacacggt	taggagtgc	agggtagaat	tcatgtttta	ttcatctttg	420	
ggtctgtagc	accagcaaa	gtgctcagta	aatgcgcagt	aattgatttg	acctctgaac	480	
aaatacacac	tgtactaaga	atctacacac	cgaagacaa	aaacaagaca	aatttgagtg	540	
ctacagggtg	cacgcttggc	atcacacatg	tgcctgtgta	ttcctctagg	tggttaccag	600	
gagctctgcc	actgcatgtc	cactagtgc	gggttcgctc	caccacccca	gctgggtagc	660	
cgctgctctc	acataagggg	tcgaattaaa	attgccagga	ataaattccc	ccggactttg	720	
acttctcaag	agctaagaag	gtttgctgag	tattctggca	tgatgtttgg	tgatcaaaca	780	
actgctggcc	aaaaatgatg	agtatttccc	cctcttgctg	aagatgtgct	ccatacaata	840	
gtccatcaca	ttcatcattc	atcagtctgg	aagtgtgcag	aacaacatgt	aatagataat	900	
atgattggct	gcacacttcc	agactgatga	atgatgaatg	tgatggacta	ttgtatggag	960	
cacatcttca	gcaagagggg	gaaatactca	tcattttata	tattacatgt	tgttctgggt	1020	
tttttttttt	tcgaatgtcc	agcctaaact	ataaagtact	ttgagaacgc	acagtgcgcc	1080	
ataagcttgc	caataaagag	tcctctgtgg	tatggaactg	gcttatttca	tacacaatct	1140	
gcaaacaatg	agggcactat	tggaaacata	ctgtgctgca	cagagcattt	acaccgctta	1200	



tctttaatct tccccagcaa tcccttgcttt gtgogcattt atgataccttg ctctcagaag 1260  
 tccacatact tttccccaac cgttaacaaat tatttaactc atctaatagt tgtatgtccg 1320  
 cgcagtctga aaacagtaat tgtccttggg aagaagttag ttttaagagag ctctagggca 1380  
 ctcatcacaa ctccagccct gccctccatg tggtagcagc tctttggact ggggctaagt 1440  
 gcttattctt gtgtttcatt cctggtaagc tcaatttctt taccttagga taactttgct 1500  
 ggaaaagggc tcagattcag ccgaccattg tggcctctgt ggctgtcaca gcttgtccct 1560  
 gacatgctat gatgttgggt ccccttctca tcccttggg atttcttctg ctggcccaca 1620  
 gccagaacaa ctaggccttt tactccacca tcccttgggt ttcttttgggt tegtgtgtaa 1680  
 aaatcaatcc ttctaccatc catgcatagc aatttctaaa aactgaattt caagagcagt 1740  
 atctgaagaa acaaacatga tttggctcct ttagtaaaaca gaataaattt taataaatca 1800  
 actttgaaat agttgtaaga gtttaagaaa agcacaaaac tgagatcatc agagcagctt 1860  
 ggctcaaaag gacaggcagc aggattctac agggtttgag ccttcctaag tgaagctgtt 1920  
 tctgcaggc tccctgctcc aagctcctag ctaacagccc cttctcccac gattggcaac 1980  
 aaagagcaaa aataactttg tacttgatgc tgagtcagtg taaaaagcca taaaaaattc 2040  
 cctctaaatg tcaaaatgtt tgccctcttt gaggtctctc tccctctact ggggtctggat 2100  
 aaattagcac tgggcttata ttgagtcaca gatctgggcc ctgccacaga gagcttctc 2160  
 ctagtgtgtg atgtttttt tccaaactat tgatacaaaa tgcactggaa tagaaatcaa 2220  
 cagaaactgg tcaaagggtg ggcatacaca ttctcatgta gatgtaaagc tgtgcttaga 2280  
 attcctttgt ggagtcctgg ttggtcttgg ttttcttgggt gtttgattca tttttttacg 2340  
 taaattacaa aaaccctcca catttcttca tggattgtat tagtccatgt tctccagaga 2400  
 agcagaacga gttggatgta tgttttggaa gagattatga ggaaccggct catgtgatga 2460  
 aggaggttga gaggtcctgt gctctgccat ctgcaagctg aagacctgga aagctgaggg 2520  
 tgtggctcca gtctgagtc gaaggcccaa gaaccagggg aaccaacggg tagattcca 2580  
 ggttgaaggc aggagaagat ggatgtccca gctcagcagg caggcaggaa gcaaatgggg 2640  
 taaattctc cttctccac cttttgttcc attcaggcct tcaacagatt ggatgagcgc 2700  
 cccccaccc ccacactagg gagggccatc tgctttactg agtcggctga gtcaagtgcc 2760  
 agcctcatcc caaaacactc tccagacaca cgcagaaatg tttcatctgg gcacctgtg 2820  
 gccagtcatg ctgacacaca gaactaacca tgacatggat tcttcttaaa gcagtgatag 2880  
 gagcgaacag aaacattttc ataattttca attattttta atgaaaacta tatctgatgg 2940  
 aattgtttaa acctagtctg gccacacatt atttcttggg accgcccctc cttcaatccc 3000  
 ttggacactg atgactttat gccagatta cactggaggc ctgtgctgat tttctaacac 3060  
 atacctgcaa ctgagctggc aaaaagaaaa ctaggcaagt atgacagata catgatgcac 3120  
 aggctaagt caaaggaaag aaaaacacca actgcaggga tgagggactc acccctttag 3180  
 aagtttctac ttgagcagct agaagactac aatgccactc atcaaaacag tgactcaggg 3240  
 ggagtatttg ggataaagga ggaatctgat gttggaggtc aaatttgaag tgtctttaag 3300  
 acctacaggt aacgagacag ctggacaaac acatggaact caggacaaag gctctaagga 3360  
 cagcacagca gctgacatcc tgtgtgacag ccttgaaagc agcaggcccg ccgctcacat 3420  
 ttggaagggt aaaatgggta caatgttgtc tgccactttg gggccttctt gggtcacatg 3480  
 cattttacat ttatgcagtt gatataattt tgtttcttgg gtcttttata cattagacac 3540  
 catgattctc aatcctttgt tattttgtat taaaaaagc tgaattatta tttcaaatat 3600  
 gggcaaatta gagccttcca tattgccaag gtgtatcaac cactctgata ycaygatctc 3660  
 tcttttgaat tagttttcca gttcacacct accatttatt tcatgattgg tttcagactt 3720  
 gttcctctg gaaacactcc ctaacaagca cccttgagg aatgaagaca caccacacac 3780  
 atctacccca ttactgcatg tactcaagag tcagctttta tatgatctct cccaagtgtc 3840  
 cctataatgg ggatctttca ctcaccctaa agtgaggaca aaataactga aagcatgagc 3900  
 ccagtgcctg taggtgtgca attaacctca gaccaaggaa gtgccgaacg catctggctt 3960  
 ttagcaaggc acctgacaaa gtccttcagg atgtttttgt acatgagcta gagaaatgta 4020  
 cctggagaac agcttctact gccagatgat cttactcaaa agatgcagat taagcaaaat 4080  
 atcaacccaa aggggtggtcc ctgatggccc accagcccct gtgcctggct cgtttcctat 4140  
 gtttcttaga tttggtttca gacttgctcc tctgcagac actccctaac cagcatcctt 4200  
 gcagaaaact ggtgaactag aaaaggcctg tgtgggtcac gtggccaccc aacaccacag 4260



```
<210> 570
<211> 951
<212> DNA
<213> Homo sapiens
```

```
<210> 571
<211> 819
<212> DNA
<213> Homo sapiens
```

<400> 571						
cagcttaaaaa	atggttttctt	gaaatcagtg	attagcattc	actcaccagt	acctacta	60
aggggtaggc	actggtttgt	actcctggga	atacaggagt	acaccagaat	ttattttctgc	120
ttattgcttt	tgttgcaaat	gccgtggctt	catctgagga	attctagaat	tcagaggggtg	180
tagccctcca	ctctgctgtc	ttgctatctg	ctctcattgc	atccgtttaa	cctgcattct	240
gaaagatgtt	tctcaggttt	ttccttgacg	atthttcttct	tttctgattc	tgacaatgtt	300
ttaaatcatt	gtactgtggt	tatcatttct	ctgcatttat	tttaccctac	ttcctttgtta	360
acttgtccta	ttgtctttta	atthttctgct	gttcttttatg	gctttcaact	tcataaataa	420
catgttttct	caaattctctt	tgtgaattcc	agagagggcc	aggcacgggtg	gctcacatct	480
gtaatcccag	cactttgggg	aggctgagac	gggtggatca	cttgagggtca	ggagtttgag	540
accagcctgg	ccaacatggt	gaaatcccg	ttcactaaaa	atacaaaaat	taccaggga	600
tggtggcggg	cgctgttaat	cccagggtact	cgggaggctg	agggaggaga	atcgcttgaa	660



```
<210> 572
<211> 203
<212> DNA
<213> Homo sapiens
```

```
<210> 573
<211> 132
<212> PRT
<213> Homo sapiens
```

```
<210> 574
<211> 63
```



<212> PRT

<213> Homo sapiens

<400> 574

Met Thr His Ser Ser Ala Trp Leu Glu Arg Pro Gln Glu Thr Tyr Asn  
5 10 15

His Gly Gly Arg Arg Arg Gly Ser Lys Ala Arg Leu Thr Trp Trp Gln  
20 25 30

Glu Arg Thr Ser Glu Gly Gly Asp Cys His Lys Leu Phe Phe Phe Glu  
35 40 45

Thr Arg Val Trp Pro Cys Cys Pro Gly Trp Ser Ala Val Ala  
50 55 60

<210> 575

<211> 77

<212> PRT

<213> Homo sapiens

<400> 575

Met Val Lys Ser Arg Phe Thr Lys Asn Thr Lys Ile Thr Gln Ala Trp  
5 10 15

Trp Arg Ala Pro Val Ile Pro Gly Thr Arg Glu Ala Glu Gly Gly Glu  
20 25 30

Ser Leu Glu Pro Gly Arg Leu Arg Glu Glu Asn Arg Leu Asn Pro Gly  
35 40 45

Gly Arg Gly Cys Ser Glu Pro Arg Ser Cys Cys Cys Thr Pro Ala Trp  
50 55 60

Ser Thr Glu Gln Asp Ser Ala Ser Lys Thr Asn Lys  
65 70 75

<210> 576

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> unsure

<222> (42)

<223> Xaa = Any Amino Acid

<400> 576

Met Leu Gly Lys Ser Arg Ala Val Cys Leu Pro Ser Thr Thr Val Thr



```

<210> 577
<211> 58
<212> PRT
<213> Homo sapiens

<400> 577
Met Tyr Leu Glu Asn Ser Phe Tyr Cys Gln Met Ile Leu Leu Lys Arg
          5                      10                      15

Cys Arg Leu Ser Lys Ile Ser Thr Gln Arg Val Val Pro Asp Gly Pro
          20                      25                      30

Pro Ala Pro Val Pro Gly Ser Phe Pro Met Phe Pro Arg Phe Gly Phe
          35                      40                      45

Arg Leu Ala Pro Pro Ala Asp Thr Pro
          50                      55

```

```

<210> 578
<211> 52
<212> PRT
<213> Homo sapiens

<400> 578
Met Gln Leu Ile Tyr Leu Cys Phe Leu Gly Leu Leu Tyr Ile Arg His
          5                      10                      15

His Asp Ser Gln Ser Phe Val Ile Leu Tyr Tyr Lys Lys Leu Asn Tyr
          20                      25                      30

Tyr Phe Lys Tyr Gly Gln Ile Arg Ala Phe His Ile Ala Lys Val Tyr
          35                      40                      45

Gln Pro His
          50

```



<210> 579  
 <211> 57  
 <212> PRT  
 <213> Homo sapiens

<400> 579  
 Met His Phe Thr Phe Met Gln Leu Ile Tyr Leu Cys Phe Leu Gly Leu  
                     5                    10                    15  
 Leu Tyr Ile Arg His His Asp Ser Gln Ser Phe Val Ile Leu Tyr Tyr  
                     20                    25                    30  
 Lys Lys Leu Asn Tyr Tyr Phe Lys Tyr Gly Gln Ile Arg Ala Phe His  
                     35                    40                    45  
 Ile Ala Lys Val Tyr Gln Pro His  
                     50                    55

<210> 580  
 <211> 68  
 <212> PRT  
 <213> Homo sapiens

<400> 580  
 Met Glu Leu Arg Thr Lys Ala Leu Arg Thr Ala Gln Gln Leu Thr Ser  
                     5                    10                    15  
 Cys Val Thr Ala Leu Lys Ala Ala Gly Pro Pro Leu Thr Phe Trp Lys  
                     20                    25                    30  
 Gly Lys Trp Val Gln Cys Cys Leu Pro Leu Trp Gly Leu Leu Gly Ser  
                     35                    40                    45  
 His Ala Phe Tyr Ile Tyr Ala Val Asp Ile Phe Met Phe Pro Gly Ser  
                     50                    55                    60  
 Phe Ile His  
                     65

<210> 581  
 <211> 78  
 <212> PRT  
 <213> Homo sapiens

<400> 581  
 Met Leu Glu Val Lys Phe Glu Val Ser Leu Arg Pro Thr Gly Asn Glu  
                     5                    10                    15



Met Met Phe Gly Asp Gln Thr Thr Ala Gly Gln Lys  
50 55 60



<211> 77

&lt;212&gt; PRT

<213> Homo sapiens

<400> 584

Met Cys Leu Cys Ile Pro Leu Gly Gly Tyr Gln Glu Leu Cys His Cys  
5 10 15

Met Ser Thr Ser Asp Gly Phe Ala Pro Pro Pro Gln Leu Gly Ser Arg  
20 25 30

Cys Ser His Ile Arg Gly Pro Ile Lys Ile Ala Arg Asn Lys Phe Pro  
35 40 45

Arg Thr Leu Thr Ser Gln Glu Leu Arg Arg Phe Ala Glu Tyr Ser Gly  
50 55 60

Met Met Phe Gly Asp Gln Thr Thr Ala Gly Gln Lys  
65 70 75

<210> 585

<211> 51

&lt;212&gt; PRT

<213> Homo sapiens

<400> 585

Met Val Tyr Arg Phe Gly Gln Met Ser Asp Asn Pro Phe Tyr Ile Leu  
5 10 15

Ala Ser Leu Gly Ser Ser Ser Cys Arg Asn Gly Leu Ala Ser Lys Trp  
20 25 30

Arg Gln Ala Asp Pro Ser Asp Gly Tyr Met Glu Pro Cys Phe Gln Leu  
35 40 45

Leu Phe  
50

<210> 586

<211> 61

<212> PRT

<213> Homo sapiens

<400> 586

Met Leu Val His Ile Tyr Ser Cys Cys Gly Met Val Tyr Arg Phe Gly  
5 10 15







Ile

<211> 157

<212> PRT

<213> Homo sapiens

<400> 589

Ser Val Thr Cys Asp Arg Leu His Ala Asn Ser Arg Val Arg Tyr Leu  
20 25 30

Met Glu Ser Met Lys Ala Leu Glu Lys Leu Val Lys Arg Arg His Pro  
50 55 60

Met Ser Gly Val Cys Val Ile Leu Thr Val Leu Lys Pro Thr Ser Ile  
85 90 95

Lys Lys His Arg Val Arg Asn Arg Arg Lys Leu Lys Ser Cys Leu Trp  
115 120 125

Val Asp Val Lys Ile Thr Gln Leu Gln Leu Leu Ser Leu Lys Met Gly  
130 135 140

Ile Met Gln Glu Gln Ile Met Gln Arg Met Leu Thr Asn  
145                      150                      155







Thr	Lys	Ala	Asn	Glu	Gln	Ala	Asp	Leu	Leu	Val	Ser	Ser	Ala	Phe	Ile
1				5					10					15	
Glu	Ala	Gln	Glu	Leu	His	Ala	Leu	Thr	His	Val	Asn	Ala	Ile	Gly	Leu
			20					25					30		



```
<210> 593
<211> 271
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc_feature  
<222> (1)...(271)  
<223> n = A,T,C or G
```

[illegible]

```
<210> 594
<211> 376
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(376)
<223> n = A,T,C or G
```

<400> 594  
cctttggggg nggggggaac ctttaccatt gtnccctttt atttcatttg gttnggggttc 60  
gcgcctcctnn gggccaacaa agttatcgtt nttgaagaga anattttttt ggnttnqnc 120







<223> n = A, T, C or G

tttcnatacc	caaaantacc	ccatattang	accanacatt	tgtctnggaa	aaattaccat	60
tntntaacnt	ttggggccacc	tgagannaaa	tggtgtgaat	ncatgataag	atggancagn	120
attnctctta	agatnngatn	agaccccggt	tttcacggaa	catatccaag	nacccaatag	180
gnaacaagcc	acggngngag	tcacaaacat	atattcttta	ctctcataat	cogtnncaca	240
naactnttgn	acttgac					257

<211> 222

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

nntggntacc	gtcnaaactt	nncttggtac	ccgagctcgg	atccactagt	ccagtgtggt	60
ggaattccat	tgtgttgggc	tataagctgt	aatagtggag	ncgtgctnng	ttcattgcan	120
nagnccctcc	gcanncaacn	ttgnnacaac	ctgtgagnag	gcataaaatt	attcacataa	180
tcactactgc	atgaanctga	ctcaaacqca	tcacntaca	cc		222

<211> 238

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

gcatgacatc	ancgatgtnt	ttggnnacct	ganattngct	aaaactngng	natgccgggn	60
atgnaggttt	ggtantgata	tatgcactca	catctcatgg	ggacgtttca	tgtggagtgn	120
tcgacaangt	tgctgnancn	gagaagtgat	gatctcagtt	gaaaggggtca	tgtgaataca	180
cnttacactt	gaaaaaqaag	cacattqgga	atatcacgaa	acgncaccca	acatcctg	238

<211> 232

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G



```
<210> 601
<211> 547
<212> DNA
<213> Homo sapien
```

[illegible]

```
<220>  
<221> misc_feature  
<222> (1)...(826)  
<223> n = A,T,C or G
```

[illegible]



aatcaagatc	tttaggccag	aaatcatgaa	nanttttana	attattttan	gaatctgtgg	720
cttctcttct	taaaatngaa	aaaaaaattg	tttaaaccce	naaggtctga	atacccaagc	780
nccttgaach	anagaacaan	gccggagcac	cccttcccaa	atcccc		826

&lt;210&gt; 603

&lt;211&gt; 817

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(817)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 603

nnangacttt	tgtggtntta	tacaattntt	ttttctattt	ctatgaagag	aaagccacag	60
agtcctaaaa	taattctaaa	actcatcatg	actttcttgc	ctaaaagatc	ttgatttcaa	120
tcgtgcctag	ttttgcttta	atcacttgct	tgagaaatac	ataaatcccc	acttaagatt	180
agtgacggca	tatctctggc	acccatttct	ggttctatta	aaattcctag	agatgtcaaa	240
aattacatta	ggccacctga	caggctatac	ctagaagaga	aaaaatgatt	tgtaaaagca	300
gtggggctat	ttgcgattgc	tttttttttt	tcttaaatac	cacctattag	gttgaaaacc	360
tgaaattgca	gctttctgta	gaaatggcgg	aagacaaact	aacattttta	aagcgctctc	420
atthagctct	gatgagtact	acaccctga	tattcttctg	atactaaaat	aattttccta	480
gtgtagtcta	aactttttta	aaaagacatg	taatccgagg	agtttgtaac	tcaaaacgag	540
tgcacttagg	aggtatcgca	agccgtttct	ggattaaatt	cccagctagc	ttgcttgctt	600
agcagggggc	gnaaanaaag	acatctgcag	cctagggaag	aaaacctttc	gcattgttct	660
tacgtgttta	cggtatttta	tttcttanaa	caaggcngaa	ttgggactcg	aatgggttcag	720
ttgggggtgg	ggatcccctg	gtncataaaa	ngtcanaaag	anggtacagg	cggaacncca	780
agggtcgtcc	tgcatttana	ctcgggaattt	tggtgcc			817

&lt;210&gt; 604

&lt;211&gt; 694

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(694)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 604

cttttcaaat	catttttnct	cttctaggta	tancctgtca	ggtggcctaa	tgtaattttt	60
gacatctcta	ngaattttta	tagaaccaga	aatgggtgcc	agagatatgc	ctgcactaat	120
cttaagtggg	gatttatgta	tttctcaagc	aagtgattaa	agcaaaacta	ggcacgattg	180
aaatcaagat	cttttaggca	anaaagtcac	gatgagtttt	agaattattt	taggactctg	240
tggctttctc	ttcatagaaa	tagaaaaaaa	aattgtataa	aaccacaaaa	ggtcctgaat	300
agccaaagca	acactganca	aaaagaacan	agcaggggaag	caacacacta	ccngaattca	360
aattatacta	ccagggtgta	gtaaccacaaa	cagcattcta	ttggcataaa	atagacacca	420
agaccaatgg	ancagaataa	agaaccccac	aaataaatcc	atataatntac	cgccanctga	480
ttatcaataa	cnaacaccaa	gaacatatnt	taagggacnt	nctattcaat	aantagtgtc	540
ggnaaaaact	gggaaatcca	tatgcagaaa	naatgaaact	agacccctat	ccctcaccat	600



<210>	605
<211>	678
<212>	DNA
<213>	Homo sapien

<400>	605								
taaaaatcta	gactacacta	ggaaattatt	ttantatcag	aagaatatca	ggggtgtagt				60
actcatcana	gctaaatgag	agcgctttaa	aatgttagt	ttgtcttcgc	ccattttctac				120
agaaaagtgc	aatttcagg	tttcaacct	ataggtgata	tttaagaaaa	aaaaaaaaagca				180
atcgcaaata	gcccccactg	ttttacaaat	cattttttct	cttctaggta	tagcctgtca				240
ggtygcctaa	tgtaattttt	gacatctcta	ggaatttttaa	tagaacacaga	aatgggtgcc				300
agagatatgc	ctgcactaat	cttaagtggg	gatattatgta	tttctcaagc	aagtgattaa				360
agcaaaaact	ggcacgattg	aatcaanat	cttttaggca	agaaagtcac	gatgagtttt				420
anaattattt	taggactctg	tggctttctc	ttcatagaaa	tagaaaaaaaa	aaattgtata				480
aaaaccacaa	aaggctcctga	atagcccaa	gcaacactga	acaaaangaa	caaagcagga				540
agcaacacac	taccggaatt	caattatact	accaaggtgt	antaaccaa	acagcattct				600
attgggcata	aaatagacca	aagaccagt	ggaaacagaa	taaagaancc	caaaataaat				660
cctatatatta	cngccccnc								678

<210>	606
<211>	263
<212>	DNA
<213>	Homo sapien

```
<220>
<221> misc_feature
<222> (1)...(263)
<223> n = A,T,C or G
```

<400>	606						
gtggggtcng	cancagccaa	ctcagcttcc	tttcgggctt	tgttagcaga	cggatcatcc		60
tctagtccac	tgtgntcaaa	ttccattgtg	tgggggcnc	tcgcctcggc	canagatctg		120
agtgancana	cntgtcccca	ctgaggtgcc	ccacagcngn	ttgtnttcag	cangggctna		180
caactcgacc	ggcagcgnan	ggctggcaga	antgngcgcc	tnnctcattc	ctacgcngtn		240
ngccgcagga	aggangacag	gcc					263

<210>	607
<211>	22
<212>	DNA
<213>	Artificial Sequence

<220>  
<223> Primer



<400> 607	
ccatgtgggt cccggttgtc tt	22
<210> 608	
<211> 22	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer	
<400> 608	
gataggggtg ctcaggggtt gg	22
<210> 609	
<211> 40	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer	
<400> 609	
gctggacagg gggcaaaagc tggggcagtg aaccatgtgc	40
<210> 610	
<211> 27	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer	
<400> 610	
ccttgtccag atagcccagt agctgac	27
<210> 611	
<211> 46	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Primer	
<400> 611	
gatagagaaa accgtccagg ccagtattgt gggaggctgg gagtgc	46
<210> 612	
<211> 40	
<212> DNA	



caagagtccttg aggccgacca agagccaggg agccagatgg tggaggccag cctctccgta 240



```
<210> 617
<211> 449
<212> PRT
<213> Homo sapien
```

Met	His	His	His	His	His	His	Ile	Ile	Asn	Gly	Glu	Asp	Cys	Ser	Pro
1				5					10					15	
His	Ser	Gln	Pro	Trp	Gln	Ala	Ala	Leu	Val	Met	Glu	Asn	Glu	Leu	Phe
			20					25					30		
Cys	Ser	Gly	Val	Leu	Val	His	Pro	Gln	Trp	Val	Leu	Ser	Ala	Ala	His
		35					40					45			
Cys	Phe	Gln	Asn	Ser	Tyr	Thr	Ile	Gly	Leu	Gly	Leu	His	Ser	Leu	Glu
	50					55					60				
Ala	Asp	Gln	Glu	Pro	Gly	Ser	Gln	Met	Val	Glu	Ala	Ser	Leu	Ser	Val
65					70					75					80
Arg	His	Pro	Glu	Tyr	Asn	Arg	Pro	Leu	Leu	Ala	Asn	Asp	Leu	Met	Leu
				85				90					95		
Ile	Lys	Leu	Asp	Glu	Ser	Val	Ser	Glu	Ser	Asp	Thr	Ile	Arg	Ser	Ile
		100					105						110		
Ser	Ile	Ala	Ser	Gln	Cys	Pro	Thr	Ala	Gly	Asn	Ser	Cys	Leu	Val	Ser
		115					120					125			
Gly	Trp	Gly	Leu	Leu	Ala	Asn	Gly	Arg	Met	Pro	Thr	Val	Leu	Gln	Cys
	130					135					140				
Val	Asn	Val	Ser	Val	Val	Ser	Glu	Glu	Val	Cys	Ser	Lys	Leu	Tyr	Asp
145					150					155					160
Pro	Leu	Tyr	His	Pro	Ser	Met	Phe	Cys	Ala	Gly	Gly	Gly	Gln	Asp	Gln
				165				170					175		
Lys	Asp	Ser	Cys	Asn	Gly	Asp	Ser	Gly	Gly	Pro	Leu	Ile	Cys	Asn	Gly
			180					185					190		
Tyr	Leu	Gln	Gly	Leu	Val	Ser	Phe	Gly	Lys	Ala	Pro	Cys	Gly	Gln	Val



	195		200		205										
Gly Val	Pro Gly Val Tyr Thr	Asn Leu Cys Lys Phe Thr	Glu Trp Ile												
210		215		220											
Glu Lys	Thr Val Gln Ala Ser Ile Val Gly Gly	Trp Glu Cys Glu Lys													
225		230		235											240
His Ser	Gln Pro Trp Gln Val Leu Val Ala Ser	Arg Gly Arg Ala Val													
	245		250												255
Cys Gly	Gly Val Leu Val His Pro Gln Trp Val	Leu Thr Ala Ala His													
	260		265												270
Cys Ile	Arg Asn Lys Ser Val Ile Leu Leu Gly	Arg His Ser Leu Phe													
	275		280												285
His Pro	Glu Asp Thr Gly Gln Val Phe Gln Val	Ser His Ser Phe Pro													
	290		295												300
His Pro	Leu Tyr Asp Met Ser Leu Leu Lys Asn	Arg Phe Leu Arg Pro													
305		310		315											320
Gly Asp	Asp Ser Ser His Asp Leu Met Leu Leu	Arg Leu Ser Glu Pro													
	325		330												335
Ala Glu	Leu Thr Asp Ala Val Lys Val Met Asp	Leu Pro Thr Gln Glu													
	340		345												350
Pro Ala	Leu Gly Thr Thr Cys Tyr Ala Ser Gly	Trp Gly Ser Ile Glu													
	355		360												365
Pro Glu	Glu Phe Leu Thr Pro Lys Lys Leu Gln	Cys Val Asp Leu His													
	370		375												380
Val Ile	Ser Asn Asp Val Cys Ala Gln Val His	Pro Gln Lys Val Thr													
385		390		395											400
Lys Phe	Met Leu Cys Ala Gly Arg Trp Thr Gly	Gly Lys Ser Trp Gly													
	405		410												415
Ser Glu	Pro Cys Ala Leu Pro Glu Arg Pro Ser	Leu Tyr Thr Lys Val													
	420		425												430
Val His	Tyr Arg Lys Trp Ile Lys Asp Thr Ile	Val Ala Asn Pro Glu													
	435		440												445
Phe															

&lt;210&gt; 618

&lt;211&gt; 385

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(385)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 618

ctgtgctgag	aaccaaagc	tatgancact	gcttttccaa	atgtccataa	naccaacatt	60
tttatcacta	ccaccatcac	ctgggagctc	nttagaaagc	tagtctcccg	ggcaccaccc	120
tggcctactg	aacctaattgt	gcatttaaca	agattnacgt	ngaaatctgc	aaagcacagg	180
ggcngataac	agtaccacct	gntctgggtc	ctanccccan	gacccttaca	gtctaactgg	240
gacacaaggg	cttnaaatca	aattgcctat	cattaagata	tacaanganc	ntgagaaact	300
gctncactta	tntattaagg	ngctctaaga	cttagaaacn	aaangcantg	ctgagangat	360



385

<211> 869

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (869)$ 

<223> n = A,T,C or G

<400> 619

gatatcccg	gaattcgcg	cgcgctcgac	ctctacttgt	ttagacataa	atgcagtcta	60
gcattaaaga	tcctttaaaa	aaatgttttc	ccaatggtta	aaagacaagc	tcaaataaat	120
gaactctcat	acatatgcca	aaattgatga	gtagataaat	atttcagtag	gtagttacta	180
gctttctgtg	tatgagtaaa	catatgggag	aaatttaaaa	cactaaagta	gactcaatga	240
aagcatagta	tcctatgtat	tcgtttttca	gaaatgtcta	atgaagggaag	gaaacaatga	300
atgaatgcc	ttattcctct	tagagtgtctg	ggacatggtt	ttgctgaaa	acttcatgtg	360
aattttatat	tttgctacac	attacaccca	tcttagactt	atcgtataa	gacataaggc	420
atatcttatg	tcttacatgt	ataataatct	aagcagaaca	aaaaataacg	aaatattttc	480
ttcccaaat	ttttgagaca	gatggatttt	cgggaaagat	gtgtttagct	tttaatcctg	540
tggttttgtg	taccacctgg	cacactagag	tgttgctcta	attcagtgag	ttgtaactct	600
gggtgaacag	tggaaatact	agggtagatt	ttaaaaatgc	taatgctcgg	gcctcgctga	660
agaccaaaatt	aattggaatc	tctgngggng	gnattgatct	ttttataatc	tttctanang	720
attctaattgg	gcttccaggg	atgaaaaccn	ctgntggagc	tnggaacctt	ccttttagttt	780
ggagaaaacc	cgatgaggg	ntnttaggcn	cgcctnttt	ttggcctggg	cttcccccc	840
tatnntnttt	tggaaagnc	cnaattttt				869

$\langle 210 \rangle$  620

<211> 339

<212> DNA

<213> Homo sapien

 $\langle 220 \rangle$ 

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (339)$ 

<223> n = A, T, C or G

<400> 620

gngcgggcct	cnccggtgctt	gctctcgtcg	ccgacgctct	ttttccacca	gctgtaggan	60
aagcccgaa	accactggtc	ccccgggtag	ccaagtacc	actggtcctc	ctggctcctg	120
acgctncggg	tcttcctcgt	ggcgtagact	gccagcttcg	gagaccctc	agccctccc	180
cgctttctc	caccccgagg	ggccatcagt	agcgagctac	tgcttcggcc	acaacctccc	240
agcangatag	cccgcggttt	ccaatctgcg	aaaggaggac	cgccnagccc	gaaatgccna	300
qcccagcnat	cactgccacg	ccgaqccnag	cgctcgtgc			339

 $\langle 210 \rangle$  621

<211> 267

<212> DNA

<213> Homo sapien



<400> 623



```

aaaactgtac tcgcgcgctg catgtcgaca ctagtggatc caaagaatcg gcacgagcga      60
aaangctcan gcagcccggc tggcgcgcgc cgctcctccc cccaggaaaag ccaangtgga      120
ngctgatgtg gctgcangag ctcgtttcac agccccctcan gtgganctgg ttggggccgcg      180
gctgccangg gcggaagtgg gtgtccccc cangtcagccc caaggctgcc cctcaciaaag      240
cactggtggt ttgcctccac tgccaccttg ggtccgaac ccgctccctt gctgtggang      300
cccaccgtgg gaatccaggt ccccagggtg actgcctgcc ttgccctcac tgcccactct      360
gccacacttt ccctgcctag anaccgggaa ggggctgtgt cgggtantgg gccacactgg      420
atgtggcagc accgactgtg ggggtggacc tggccttgcc ggggtgcaaaa gtggggggccc      480
ngggaaaagc acctgaagtg gccctgaaaa atccccctt aatttttccc caatttgggg      540
ctcnaacaaa aggaaattgc tgaagccaan ggtaccaagg tcacccttaa ggccagggtg      600
aaaagggtccc aaaattccaa tccccacnt ttgggcttnc ctcttggaac cccggccccc      660
tctcntgaan ttttaaaaaa n                                           681

```

<210> 624

<211> 661

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(661)

<223> n = A,T,C or G

<400> 624

```

attggtctta ctgtaccacc ggggtggaaat cgatggccgc ggcgtctaaa tatccgattt      60
tttttttttt tcctcttctg actgtccatg gacaaatgaa actaacttaa tctaactaaa      120
aaacacaact atattttgaa gattttctat ctgcactcaa ggacactttc cacnccggtg      180
ttgttacctt ttggtcttgt ctctgaacat gaaattnatc tcaagggatt ngatttctgg      240
acctcctatt cctgctatgg gtttgatatt tcttgggctc cagggccact gttgcattgg      300
gntgacagnt acctcctagc ccatanctc ctatcttggg aaacaaacct aacaactacg      360
tgtaccttcc atagatctct gattgagtct cagtatnccg ttgctcatgg gcgattcact      420
tgaatccgtn attggtgcca acaatcctga ctcatggggn aatggatcct atcacgttcc      480
cctgattngc aacccctgta tacatanatc taatcgcata gaatctagcn tnggntatgc      540
gcggctacgc tatcagggnt tgntaactat ngcatggcta cgaancctga tcatgatcna      600
gggtcatgga ctcttatcag ggggggttggg ccgngcttct ttttcnnacc ttggtaaaac      660
c                                           661

```

<210> 625

<211> 181

<212> DNA

<213> Homo sapien

<400> 625

```

gcaacaatca gatcatgtta aagtaaactc ccattgccct ggatcacttc aggatttaat      60
tgtccaagga gagcagggtt ctctgtgtaa aaaaagggtg ggaaatgttt gagagtaaaa      120
aatacaaaat tcaaccggtc gaaaatacac cactccattc agtgctctac ccccataagc      180
c                                           181

```

<210> 626

<211> 181

<212> DNA



tttggngn	ggtgtctcnt	ttgggtggac	tttttgggtc	gtaggccccc	aaggccgtta	60
atcccgtaat	aacggaagac	gaagaagagt	cagaagagtg	cttcataaag	gatcgggacg	120
agactacctt	agaggaataa	aggaaaaaag	cagaggagga	agagtggtag	aaggagtcag	180
aagaaaccca	cacgtcgttc	tgaacctgga	gccttatcaa	aaaggcttag	ataaacgata	240
gcgatctcga	tatcgagctc	aagaggtagg	tttagagact	tctcgtcctc	gagagcgaaa	300
tggaagatct	cgacgacgat	aagaagttaa	agtgtagagg	gtgcttgagg	agcgcgtgga	360



```

aggattctgc ggagggaccc atcgacgtag agacttgaag gcctactaag gtccacaaga      420
agcccggtct tttctccgaa tggtcggagc gtacagtatg cgacgtcgat cggcagacaa      480
gctggcggtg gactcgaagt gttcgggcga atcgacttat aatagtcgcg cgctagtaac      540
gtaggaacac gaagagtagt cgaaagaaaa cgttttagtga gggaaaagat tagggaaaaa      600
ggagaggcctt aataactaag acacttggag cctaggccaa cgcgaa                      646

```

<210> 629

<211> 617

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(617)

<223> n = A,T,C or G

<400> 629

```

gccccnccc ccctcctngg gcttatnggg acagaccac gtagtactct aaatcttctc      60
ctacgccgga caacggaccc tataccaatt cgaatcttgg aactccgac cgccggattc      120
tcttccccct tcggttccc ctttctgtcg gtacccctcc ctagtctct cctacacctt      180
cgtaccgtcg atatatagtc gccgcggact agcctattta ggtgtcctag actcgttatt      240
gatccactca ttagtctagt actatgcgtc acgtatctta gttgcctaag agggagatta      300
aatcctccac aagttccgac gaattcctgg actctcgtac tagcaaactt tcttatgagg      360
cttccttgta tatcttctgg atgtttctcg tgtcccggtc ctccgctact actagagctc      420
cttgccctat ctctagaagt agaggactct cgggttcggt ctccaaatct agcgctagag      480
ctatcgctac ccgctcgatt cccccagcgg aatcttgaaa cctgaggtag tacacaaacc      540
ctcncatct tccctcggtt gctccttctt ctcatcccc cttcccgctt tctcgggan      600
gaatctactt tancttc                      617

```

<210> 630

<211> 644

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(644)

<223> n = A,T,C or G

<400> 630

```

cnntcggcnt gggttttntt ctgagnnncc ccccccccc cccccccaaa cttacacca      60
ccaaacactt tccgccccct acctaggaga cattagaagg gtttaggctt cggcgtatag      120
taaagtcctc tacctcgga gtagagaatt cggtatTTaa attcagggtt agaggctcgc      180
tcgttagatt tatagttag gtttagaatc ggaaaccttc gatcttcctt agaagggtaa      240
taagtgaggc cctaaatccg tctaaccaag gcgttaagggt ccgtacctaa acctagtctt      300
atcttctatc aggcgcacca atataggtag gttctacttt cgtataggcc ttaaggaata      360
gttcggtagt tatcgaaggc actcctctct aggcctaggct tttctcagtc ttagtactcc      420
gggaccgtcg tcgcanaaat atcgatggac ggtagggtatc tccgcgttac gcgtcgggct      480
agggatatag agcgaattat cggcgagagg cggtcgctan gaatcggtat caatatgntg      540
ttctttaccc tacggatatc ggcagaaaac ataaaacctt ctnaccangg ataagggtatt      600
atcggacccc taaaataaca gtaacattta gantactagt accc                      644

```



<220>



<221> misc\_feature  
 <222> (1)...(630)  
 <223> n = A,T,C or G

```
<400> 633
tccttcgggt tgggtttttt ttctgaccccc cccccccccc cccctcgga aggcctctag      60
gtccccaccc gtctctctaa ttctcaggaa ccgatccacc caaccaactt actaatgtcc      120
tacagtaaac acccgagaat ataaacccac acctaggcct ccaatcctac cagggaagca      180
agaagccgta gtctagcgta ttacgaaccc gagatagaga cggagatact tagttttatt      240
ctctcggaat aggaaagacg actggggagg gaatataggg tagcgcgggg ataggggcta      300
tggcggatat gggggcgggg cgtctcttta ttctttctata ccacgtcaat aggaatgtag      360
atatacctag atgttcccgt agaaagagac gttagaggtc tccgaagcta taaaggagag      420
gcgcgaagaa acttcgtact ctagctttat ataggtagtc gctctagtcc cataagcgac      480
gagagatcta ctagatttcg gtatcgccgt cgtatgtatt cgaaatagtc ttcttccctt      540
tttcgatctc ctctctatac tacatggnga ttatagtcnt aagatagtca ggatattagg      600
atattagtta tatgacgttc gacgggacgg                                     630
```

<210> 634  
 <211> 647  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(647)  
 <223> n = A,T,C or G

```
<400> 634
ccntcggtt ggggtttttt ctgaccccc cccccccccc cctccactaa gancttaacc      60
caaccctata gtttactcgt ataggggaat cgaggagaaa taggaacgaa gagcgggtga      120
taaagagaaa gtactttcct ttatatgtta agagcttagc gtaatgactt tcgttatatg      180
gctagttgat tttatccggc gttatagggc ttagttctgg ttatctcggg tctaattccc      240
ttagtatgct cgggagtta acgaggtcac gggatagcgc gtaccctttc taaggttcct      300
ggaaagctat tcgttattta tcgcgattct cgaggtcgaa aggatcaagg atcttccctt      360
ttactaccct agtcgggtta gcggtcggtc aaaactagt tagtaccttt acctcctcga      420
aagttatagt cgaaacaacg tattagtoga aattatagcg gatagatcga gacggttcct      480
tctcgggttc tcagccggta atccctctat ttgggggtct tctccctctt cccctttgtc      540
ttccgcctta gcttccaagg ttctcgggaa gcgagggggt ctacttaagt cgntagcgtt      600
ccttataaac cncctacagg cagaccccc ttgtaaaggc tcgggggt                                     647
```

<210> 635  
 <211> 645  
 <212> DNA  
 <213> Homo sapien

<220>  
 <221> misc\_feature  
 <222> (1)...(645)  
 <223> n = A,T,C or G

<400> 635



```

ccttcggctt ggggtttttt ctgagccccc cccccccccc cccgaaactc gccttaccct      60
agatacccaa agaatagttc cactcaactt cgtctaagta aaactctaga acttccaaac      120
ataaaaagact tcgcgcggtt agctacacag cctacgggaa tctcacgaat cccgattcaa      180
gtcccactct cgaccacacc ccggtatcgt cgttttccca taccaatgtc gaaaaataaa      240
ataaaaatcca gtcaagcccc acggttaagcg ggggtagggc taggcgaaga ggcaggaacc      300
gttcgaggcc gggggctttc aaaatacaaa acaactactt aaagtttacc ctttctaaag      360
tcgggggcaa cggttaaagc acgcctctaa agtactactc gtttcgagaa ggggtagtca      420
tctcccgcat agagactctc gcgtatatca actcgcacgc cttctagcat tccgacggtc      480
gcccgcggtc acatatcttg cggattagct ccgagggact atagggttaa ttagtctagt      540
aaattctctt agaggatagt cggggtcgta gttaggcagt acgaggggac atggncctgcg      600
tcgtgctcta ccttgacagc atactcttat aaacatcttt ttcct                      645

```

<210> 636

<211> 643

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(643)

<223> n = A,T,C or G

<400> 636

```

ccttcggctt ggggtttttt ctgacccccc cccccccccc cctagcggaa aacaatcccc      60
accgagattt tattaatcgt aaaactcgcc ttcggtagca agtcttctct cttcccgtaa      120
cctggctccc tcctagnggc tttacgaacg tccctctctt tcttacggct cggaagtggg      180
tacggttaaa tccggaggng gggctaacga atccaaggct aactcctctt anagtttgtt      240
gtccnncnct ttagtaagga tccgtggagg gcgagtattt gncccccgcc ctttattnta      300
tagttcccta gtacgataaa gntaccggct atcctattac agcggataaa agttatttan      360
agggccgacg tcnccgctag acaggctaca gotagnggag gtaccgcctc cgactantcc      420
gttgnttccg acaaggcnagt ttcggttaac tccacaaact cctccgccga ctctanggtg      480
gggacggcag ttccnncggt tagtgtgcgt tatagagaag ggcatttgag ttggacgtta      540
cnttttaaca taggttattc cgtttagggt cttgcggggc cgtgggggta gtncnccggc      600
gcgttnntat cggcgatttt ccgcagtttc cgtttccggn tnt                      643

```

<210> 637

<211> 631

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(631)

<223> n = A,T,C or G

<400> 637

```

gggttntctc atttgggtgg actttttggg tcgtaggaac cggtatgnag gagtaggagt      60
cgctgggaag actagaagtt agctacggac gattagtgtg attccactct taataacgag      120
taatcgttta cgtcgggttg gtgtttcggg gttttggaga gtaagcgtag ttgtggagtt      180
tcgcatatag gtccccttac ttccggcgatc tcgtcttctg tcggttaggt tattattgtt      240
catccttcgc attagtagta ggggttggtcg gataaatcga tagctattct ttagaattcg      300

```



```

tagtcggaga attcgtgtac gaagtccttt aagttcttta agttcgcgag taagacgtgt      360
acggttat tttt tgcgtcgac gtaggtgtcg tttacgggag tttcgtttta ggggttttacg      420
tagaacgtta ttaagcacgg taatacgata gaggattacg cgacgtattc gtcttagaac      480
gtcgattttt cgaaggcgca tttgttatcg aaggggagtc cttggagaat cgagatattc      540
caagaatatt acggagatta cagatcggaa ggctcccagag atcggacgta ttaccggtct      600
cgcccgaaac gagtaggtat cntccggata a                                     631

```

<210> 638

<211> 606

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(606)

<223> n = A,T,C or G

<400> 638

```

ccccccccc ctcaaccatc nattccccac ctcaacgcga attacggttt cgaaagtcga      60
caataagtcc ggtcgagtag agggaatcag gggctggtan aaaggaccac gggcggaaaa      120
taccggtctc cttccgggga gcgacgtcgg ggaaagggaa gagagcggtc tagttcgtag      180
gcaaacaggt cagaaaagtt aaggttaaag gtcggagggg agaggatagc tagtacgctt      240
agttcggggc tcgggcgcag ggccaacttt ctcttttcgcg ttcctttact ctgcttacga      300
gttcaggctc cggagttccg cgccggaggt cgtcgcgacg ctaggaatgg ggactcgctc      360
agtccccggt tatccttcgg gattctatgt tttcgccgat agacggagac cgggtagtag      420
ggttccgctc taccgccact cgtcgcttg atccggcccc ctccgcttaa gggcgatgaa      480
agattaggtt ttagggctct acgggacgag gcatagggcg ggagaagggg ggaggggtcg      540
ggggtcgaag ggantaagaa atcgcantcg cgcgggggtcg gtagganccg aaatttttct      600
cnnctg                                     606

```

<210> 639

<211> 592

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(592)

<223> n = A,T,C or G

<400> 639

```

tcctcggct tgggtttttt tctgagcccc cccccccccc cccccgggaa cgagaaaaca      60
atcccaccct accgcgggga gtgggttgna cgcttagttc tagaatctc ggaatcgctc      120
tccggcgttg gtatgtccgg cgattccgag tatgccgaag tgtatcgctc cgtctagagg      180
ttggtatctg tttatcgca tgacgtatt gactcggatg ctttcgaagt agggggatag      240
gcgcatagat acgcctccgc ggtgtcctct gaagtggccg catccgtgga cgcagcgtag      300
acagctctgg tggacgataa cggcttctcg tactcctact ccggctatta tgtagagag      360
gacttgtttc tgaacggata taccattagc gaaggggtac cctccgctaa cgcaggcgctt      420
tctaacagtt cttccgggcy ctccgaattt agattgacgc ctccgcagca ttgtgggatc      480
ctcttcggtt agccctcttt ataggatttc tcttcgccc cgaaagangg ctggtcgtcc      540
ccggcangta tgtctagctc gaacgctttg ttactccttt gttttcgaaa na             592

```



```
<210> 642
<211> 645
<212> DNA
<213> Homo sapien
```



```
<210> 643
<211> 586
<212> DNA
<213> Homo sapien
```

```
<210> 644
<211> 646
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1) ... (646)
<223> n = A,T,C or G
```



```
<210> 645
<211> 654
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(654)
<223> n = A,T,C or G
```

```
<210> 646
<211> 645
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(645)
<223> n = A,T,C or G
```

<400>	646						
tccttcgggt	tgggtttttt	tctgagcccc	ccccccccc	ccccacgcc	aagtacacag		60
accacacaaa	aacaacgtca	acacaacttc	gggtatacgg	accttaagag	agaccccgta		120
gtagacccta	ccacagccat	ccaatagtca	aacaacaagg	gcgcacccaa	tccatccata		180
gagctatcaa	acaacggagg	ggaaaggaaa	gagcagggtc	aacttagcag	aqatcgaagt		240



```

cggcactaat tcctttcaag tactcgctcg gcttgtagtt cggggtaaag tccgctctca      300
aagggccaac gaggttttaa agcgaccccc gtatcgagtc ttcttcgtat tcattaaggc      360
gttaaaggta cgagacctag aagagagtag aattagccca ccaaatcgcc taaaccggca      420
aaaacgacca aaagtcaaag acccttaca atatacactt aaaacgcca ccccaaaaac      480
gcatcagta acgcacgtac ctttcccaag cttttctttt tttcactctc caaaacaaac      540
ccgaatatat agcgcaaaaa atatccgagg gagaattaga agctattacc cgaaaaaaaa      600
ncgganangg antaaatngt ggggaatana cgtttggttt ttctg                      645

```

<210> 647

<211> 753

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(753)

<223> n = A,T,C or G

<400> 647

```

accttacctg gtaccggggc cccctcgag tttttttttt tccaaataca actcagattg      60
tatacgaaaa gctgataata cattgaactt tgctgtttta atcccttgag cctttgataa     120
tgattttttt tgtgttaaca attgtagtat ataaaatcgg attcaccatc cttctgatgc     180
catattgatt agtttgattt tatggtgatg ggatcattgt gtgttaactg tattaagaag     240
aaatggattt gattgacttt gcatccattt ttatctgtgt tactttcatg ttttatataa     300
aagcatttct ggaccagaat aagttaagtg gtataatttg ctttttacac gtttatataa     360
ttgaagttag caatgtggca aaatctctaa tggaaataaa atgcttcaga atgatgacat     420
aaatctgagc tatttcttgc ctggagaaca agtggtattc ataataattt aatagcttct     480
gaggtgtttt gttcatgtga tgaaggctta tccaccttgt atcaattcat gggctctgct     540
ttgtttaatg tagtcagggt gttaatacna gacttaagag tcatcctact gtgataagtg     600
gtgagtgaag attacatgtc ttangaaaat tatactggga atatctctga cattaatggg     660
tttaaatgtt ttaaggctag gggatgatgc aatgganaan atncttccaa angtttctgg     720
ttgtttatat ttgnngaagn catnaagana ccg                      753

```

<210> 648

<211> 383

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(383)

<223> n = A,T,C or G

<400> 648

```

gatatcccg ggaatgcgg aggccttng gcttacgtgt ttaccgcgta gggcaaagcc      60
ttgncaaat cccggccagc ggagcggcga ggggtggggac tcacgggaag ttaaacagcc     120
tcgtcggcgt cctcgaggct ccaaaaccag gctctaggcg gggacgactg cagccgttat     180
ggaggccacc gcggctacgg ccgcggctga ggcctcccca ggtggagcgg tggcctggag     240
gggaatcttg atcctgggac agccacctgt caagaggagg cggagcgtca tgctctgga     300
agactggatg aatattctcc aggagcctga cgaaggcgaa gaagtctttg cagaggaaat     360
tgaatgctgt ctgatgctac aat                      383

```



```
<400> 651
cattgtgttg ggcagggtca tttctaaggc atgggctgga agctttttatt taaaacttta      60
catgtcttag aagcactctg gttgttgcta ggcagacaat ttacatctc ttgctatacc      120
agttgcatga agttcatcat gcatattggc tgtggaaaac cttaacagca tcatgtcata      180
```



```

aggtttcagt aaggttttaa tgaaatcatg tattaagcac ttagtatagt gcaccttaaa 240
tgtagcttc aaaacaatga caacctaaact aatgttgaaa gaagcttggtg tttgttaaatt 300
atgtcttatt gaaagatgtc atcaaatacct gttattttcta atcccttaaa gtctctcaat 360
gtattttctt ttgccatatc caatgacagg accttagttt aagccagtgg ttctctcaac 420
ttctaattcca gagatacctg ggtgtcccca agaccttttc agagcatcct tgatgtcaaa 480
accattttca taataatatt aaaatattat ttgctcattg tactcttatt ctctcccaaa 540
tattcagcga gttttccaga agctatataa catgtggtaa catcttatca ctctgacgat 600
taatagaata tngnnttttg gattcttgng tttaaaattt tctcactttg gggtttcta 660
atggmnacga ttaatagata tggntcccat gaccagangg ctttaaagca ntcaataatt 720
tttaagagac taagnactat cctttaaaga tngngaactc catcttaat 769

```

<210> 652

<211> 267

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(267)

<223> n = A,T,C or G

<400> 652

```

nnangccctt taaccattgn ggctccacg cnntggcggc cgctctacaa ctagnggatc 60
cgcnactcta gnanaangat tggtcttnt gggntgggccc ggncgggctg gggcgtaag 120
cggggctggg cgcgcgccgn ggttgnaacna ggcgcgcccg ccncacacn cccggagcac 180
cctcnttgcg gcctntcccc gctcaccg cgcgcgccgn tccgcttttt ccncacccan 240
agcnctnttt atctntgtct cctccgg 267

```

<210> 653

<211> 501

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(501)

<223> n = A,T,C or G

<400> 653

```

cccnttnacc cattgctgga ctccaccgcg gtggcgggccg ctctanaact agtgggatcc 60
ttncnatgag atnggcgang gaggacnnat ttgctatnct ggatggggct gantcntnta 120
gctnctctag cancagatgg gttatcgagg aagatgactc caangggcta nantcctatg 180
cncatcctaa aanncanctg ctgtnttcag agtacgagac acatcatcnc tnatgcattg 240
ntgancaaga cgggcangtg cttatcctca gcgangatgc ccttaaccan gagctcgaat 300
ggacntatca ccntanaggt acanntnccg caccacacac cngcttgcn cctgacgctg 360
gactggatcn cttaggccac caatnccccg ttnccacat ncctgggacn ctananatac 420
tcganggggg gcccgggtanc caattcgccc taatactgag ccttgntacg nacgctnact 480
ngngtccta ttanaacggt g 501

```

<210> 654

<211> 710



<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(710)  
<223> n = A,T,C or G

<400> 654  
gcgncctttan cncatgctgg gctccacgcg gtggcgggcgg ctctacacta gtggatccca 60  
acactgagtc caccacagna aaactcanca ccaggcagac ccacaaactg cagaatccag 120  
gctgcaattc acagactaat cntctagacc cacctcagta ccagatggta ccacacagct 180  
caaggnttta ggtttgcgtg gtanactcaa tctctatctt tcaccactgc cagcctgact 240  
tcagagatcc tgnngctctgg acagtccctca gtggcaggca actctcagga gcctcaggnt 300  
tttggcacat cccagnacca gccagctgcc acaggccctg acctntanc aacactgccc 360  
atgtattcca gacttctanc ataccacagt gccatgctga ttgcatctat agangctcag 420  
gtgncctca aanctgtgcc tctggcagna ngccccacgt ctctggcatg ccccaatgcc 480  
atgngtggn aacanttgact tctgggcagtg ntgggaattcc ctaccactga ncctgaccat 540  
aggngggganc ccattttttt cgagggggggg gcccggcccc caattccncc ntatagnag 600  
ncgtanttac gcgcnnctta ctnggcengt ngtttaacaa cgtcnntgan ctggggaaaa 660  
cccctgggng cnaccctaat taaacngent tgcannacat ccccttttcg 710

<210> 655  
<211> 202  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(202)  
<223> n = A,T,C or G

<400> 655  
ccccttttnc ctttcanccc ccccgttttg gngccgcgn acacctactn catccaccca 60  
cantcgacca cccgagcttt tttccgatcc cancactnat gcngattttt tctntgcntg 120  
ctngcctgc acctttgnta ggtcaagcct ggcccatctt cgacaacttc ctcatcacca 180  
acgatgaggc atactctgac ga 202

<210> 656  
<211> 308  
<212> DNA  
<213> Homo sapien

<220>  
<221> misc\_feature  
<222> (1)...(308)  
<223> n = A,T,C or G

<400> 656  
gctgntgaaa gaccacaccg aaaaactctn ctttccgact tccacatgat gatcngcatg 60  
tgggtggtgag agacttatca tgacgacatc gttccnacc atcgancn ctgcccgaagc 120



ccattcatgg	aggcctgggn	antttctgtga	ntgaentnga	cncatanacnc	tnccactgtn	180
tgctatccag	acttgnntng	aatatnttat	tggcnaaana	canttncgga	atgctgtgnt	240
tgnncattga	angatctgat	cactatgaga	gggtgaggac	nnctgctng	ctggcantnt	300
ntaaccn						308

&lt;210&gt; 657

&lt;211&gt; 696

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(696)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 657

accntttcca	caatnctggn	ctccccgogg	tggcgggcgc	gtcgaccagc	aacctcagct	60
gtgggtcttg	ttacagtaat	gagttactgt	aaggaaagtg	tgacatttcg	agcaatttga	120
tttgtttaaa	aactagagca	gtttcagggg	tttccttgta	aatctgtctt	atgtgtcttc	180
aatgttcttt	cttgaggagt	agagaaagga	attgttagga	atgatgcata	aaccatggct	240
tattttatct	cgctgccacc	cataatcaga	gcagattctt	gggactatga	ccctcatgga	300
gacatgacaa	ttgtgtgtgt	gggtgggtggg	agaaaagagc	tgggaatttt	tagggctctag	360
agggccaat	caggactatt	ttatggagct	ctgctcacca	actttaagtg	agcaccaggg	420
gtgngaaagc	gaatcttggg	ntcaaaaana	caatggnaag	gggtaagttg	gtatnctgaa	480
ctggccactt	cggactctta	tttaactggg	tattctcant	taaggaggcn	nggggtggtct	540
tggcttgtna	aggaaagcct	gtgcaatgga	atgactttaa	aaccccccat	taaaaaaaaaa	600
angntataaa	tcttggtct	taanaangaa	gcctgggttc	tnttanccca	ttttnccccc	660
gggaaggnaa	atnttcttag	gnaanggaag	ggaagg			696

&lt;210&gt; 658

&lt;211&gt; 698

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1)...(698)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 658

ctggactccc	cgcggtggcg	gccgctctag	aactagtgga	tccgtgttgg	ctcaattctc	60
aaggctgttg	ctgtgcggcc	tgttccccac	acgtgctgct	cagctcaggc	aagcaccgag	120
cttgtgttgt	ttcatgctca	gcgtggaggc	ccctcctcca	ggtcgctgct	ctgtgggggtt	180
cccatacact	caggctccta	ggaggagtcc	atthagaaag	ccagggtttt	tctcagagtc	240
ttagttcctt	gtgctgtcat	ccatttcaca	cgacttgggc	cctgctcggg	gcaacacagc	300
aagagaaaag	acagggaaaa	taagagaggg	accttgcaca	cacacgctct	ggaccacaga	360
gccctgtgcc	cagctcctct	gtcaatacag	gtggaatctc	gtgcaggatc	gcaggggtct	420
gtgatgccac	caaagagcag	gccgggacag	ggttaggaga	gaaaggagag	ggaagtgggg	480
gtttctccta	cgcactctta	tttgagagg	gaaaggcggg	tttgtattgg	ggttgctcgt	540
ctttgcaccc	acngcacagt	tgtgagacac	ccccatcctn	agatcaaagc	cccacataca	600
gcttggggaa	aaacaaaacn	aaacaaaaca	aaaacagtaa	acctccatgc	canttggttg	660



gnaagtttttn aatttncttc cccnacccan cttgcttc

698

<210> 659

<211> 750

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(750)

<223> n = A,T,C or G

<400> 659

ncaanctggn	ctccaccgcg	gtggcgggcg	ctctagacta	gtggatcctc	ctcatgggcc	60
tggatatctc	tgaacatatg	atgaacattg	cttatgaaaa	attatttgta	ngaaaattgt	120
gaggcctaag	aatgntattt	tcttttagtg	atgggtcttg	tttgcttctg	taaggnaactt	180
gtgggcactc	gtaagcttgg	atctctttta	tctaatacca	gnnttgagat	tttcttggcc	240
ccatagatga	attaaaactg	gcgtacttct	tgtttacaag	anggataagt	ctcctaggggt	300
aagtcttttg	gggtcccaag	tcaaaaagat	gagggattta	ccagttctct	aaccttggta	360
gccccagact	ccaaactttg	ccttctagtc	ccaagaggct	atcaaaaagc	aaaggccatc	420
ttccaccttc	ttttccanaa	cagcacacat	tccagacagt	acttgaaagc	aggaacctcc	480
ttatccctta	aaaacctott	ggaancatct	tccctctctt	gcttctacta	tgcttggccc	540
acctancatt	cncttttttc	tggaaaaccgg	aaaaancttn	tgacttnngt	tggctacatt	600
cagcttggcc	ccctacaatn	tggtttccat	ctgccctaan	gaaattttta	agggcacttt	660
ttttntggcc	cctgactttc	nntttttagg	gctttccccc	angctttgcc	cctttgggta	720
aaggggttat	tttccttccc	cttttggaag				750

<210> 660

<211> 849

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(849)

<223> n = A,T,C or G

<400> 660

tcggatccac	tagtccagtg	tgggtggaatt	cgcgggccgc	gtcgacgggc	agtagtggtta	60
tgcntntcta	aatgttataa	ttatttcaga	attactctgc	cagaaagtta	tgatcataca	120
tagaagagtt	tgtagctaac	tttgaaagta	gtggaaagtg	gttttcatgt	attgtttggg	180
ttaattttaat	tttgattata	tttggttttt	agttcaggta	atttttttgt	tgaaaacttc	240
aaatgacaat	ttcttcatgg	ttactaaaga	tactcatgt	ggagtagttt	cagatttttt	300
tctgaataca	tgtattactt	ttagagatgt	aaagatgtga	aattactaag	agagaaaccc	360
atgtgatttg	tttagtggat	caaaagtcgg	tagctccttt	gatcctaagt	gccactgata	420
gttaaataga	tactgaagct	atgggcaggc	tggattgata	agaaaaagg	agacagagaa	480
atgggaaatt	gggaaagaac	tgtgcaaata	ggaaaaggag	agagcaacag	aacagaatta	540
gtaccacagt	gccgaagtgc	cacctcaggt	acttccatct	cccattctct	gaagaattca	600
gtaacagttt	gcaaatggtc	aacacaatca	tttagtgatc	ctggttgata	ttttcaatac	660
tttctgggga	tttcttggct	ggnttcaaaa	gatgatgctg	atagttttat	tgccctgaa	720
ggtattctga	agnttancat	aattttattg	tcagtaaaat	atttgaataa	aagngganga	780



```

aggaaaatct ggcntcttat tttgggatnt cngcnggggg aangaggata taattnaccc 840
cggccttgg 849

```

```

<210> 661
<211> 653
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(653)
<223> n = A,T,C or G

```

```

<400> 661
aacttaagct tggtagccgag ctccggatccc tagtccagtg tggtaggaatt cgcggccgcg 60
tcgacctcca ttcgtttctt gtccctttttt ttcatttttt ctcattgttct attcacttta 120
ggtttctaag ataaatatta taaaataatt tttacttata aattattcac tgataccctg 180
tctttaacat gtgaaatgaa ttcaaaaagga atcttaatga gaaataatat actcatgatg 240
tttaatagat ttgatttcga aataataagc cctctgaagt cctaagttaa aaataaagca 300
acttgtttga taatttttca tcaagaatgt atctgagtc ctagagtaatt attagtagga 360
atattccatt atcacaatta cacagtataa gctatttagt ctaactttac caaaaaaggg 420
agctacttca acactgtgtg agacttttaa tgggtttgca ttgggtatgc actattagca 480
agataaccta ttttacagca gtgtttntta acctttccca tttatttgaa aggcagctaa 540
gatatagtag ttaatntaan gggctgatgc atttatatta catgtagana atgggagata 600
cnaaaggagg nggggggana tnttttgnat tcnnaagctt cnttgncaat taa 653

```

```

<210> 662
<211> 646
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(646)
<223> n = A,T,C or G

```

```

<400> 662
aaacttaagc ttggtaccgc agctcggatc cctagtcagc tgtggtggaa ttgcgggccg 60
cgtcgaccca gggacaggca gccagngctg gggtcaccag ggtccctctt tgggccctcc 120
aanagcaaca gtactggcaa cagctgggat ttgctgagca cagactctgc agcaggctcg 180
gttgagctct ctgtgcctgt tccttcatac catcctcacg cccatccatg agatgggtcc 240
agctgttttc agatgagaaa atggcacagg aagctggtta gtgacagtca gaaatgaatg 300
ctggcagctt antccttgga cccaccgcag tgcaggacct tgctcaacag ggatcaccct 360
tgtccgccac ctgttcatga ggccacccag ggtttgtgtg gtcatttgtc tcctttcatc 420
tgcttgctt caaccagctg ggtcattagg gctggggaac ccagaccca cacagtcctt 480
ctcccagang ccagacacan nctncgccac agnaaggact tcagtccccg aancaaatgt 540
ncctgggcgt anaaactgna gggnccccaa tccttggtgg ggtactgctt tgactggng 600
gaattcacc ctcattgnaa acctttccct nttncaccc ctaaac 646

```

```

<210> 663
<211> 650

```



<220>



```
<220>
<221> misc_feature
<222> (1)...(817)
<223> n = A,T,C or G
```



```

<400> 667
nnangacttt tgtggtntta tacaattntt ttttctattt ctatgaagag aaagccacag      60
agtcctaaaa taattctaaa actcatcatg actttcttgc ctaaaagatc ttgatttcaa      120
tcgtgcctag ttttgcttta atcaacttgc tgagaaatac ataaatcccc acttaagatt      180
agtcgaggca tatctctggc acccatttct gggtctatta aaattcctag agatgtcaaa      240
aattacatta ggccacctga caggctatac ctagaagaga aaaaatgatt tgtaaaagca      300
gtggggctat ttgcgattgc tttttttttt tcttaaatat cacctattag gttgaaaacc      360
tgaaattgca gctttctgta gaaatggcgg aagacaaact aacattttta aagcgtctct      420
atthagctct gatgagtact acacccctga tattcttctg atactaaaat aatttttcta      480
gtgtagtcta aactttttta aaaagacatg taatccgagg agtttgtaac tcaaaacgag      540
tgcactagg aggtatcgca agcgttttct ggattaaatt ccagctagc ttgcttgctt      600
agcaggggag ggnaaanaag acatctgcag ctagggaag aaaacctttc gcattgttct      660
tacgtgttta cgttatttta tttcctanaa caaggcngaa ttgggactcg aatggttcag      720
ttgggggtgg ggatcccctg gtncataaaa ngtcanaaag anggtacagg cggaacncca      780
agggtcgtcc tgcatttana ctcggaattt tgggtgcc                                817

```

<210> 668

<211> 826

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1) ... (826)

<223> n = A,T,C or G

```

<400> 668
cgggggggnt tacgtctctc tggacgcttt tattgtacca gggcgatccc agcccaactg      60
taccattcga gtccctactc ctgccttgc ctagggaaat aaaataacgt aaacacgtaa      120
gaacaatgcg aaagcgTTTT cttccctagg ctgcagattg tcttcttcac cgccccgtct      180
tagctagcta gctagctggg aattttaatcc agaaacggct tgcgatacct cctagatgca      240
ctcgTTTTga gttacaaact ccgcggatta catgtctttt taaaaaagtt tagactacac      300
tagggaaaat tatttttagta tcagaagaat atcagggggg gtagtactca tcagagctna      360
atgagagcgc tttaaaaatg ttagtttgct ttccgccatt tctacagaaa gctgcaattt      420
caggttttca ncctaatagg tgatatntaa gaaaaaaaaa acaatcgcan atagccact      480
gcttttacia atcatttttt tcttctaggt atagcctgtc aggtggccta atgtattttt      540
gacatctcta ggaattttta tagaccagaa atgggtgccca gagatatgcc tgcactaatc      600
ttaagtgggg atttatgtat ttctcaanca agtgattaaa gcaaaactag gcacgaatga      660
aatcaagatc tttaggccag aaatcatgaa nanttttana attattttan gaatctgtgg      720
cttctcttct taaaatngaa aaaaaaattg tttaaaccca naaggtctga ataccaagc      780
nccctgaacn anagaacaan gccggagcac cccctcccaa atcccc                                826

```

<210> 669

<211> 547

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1) ... (547)



<223> n = A,T,C or G

<400> 669

cattgtgttg	gggaaaaaat	gatttgtata	agcagtgggg	ctatttgcga	ttgctttttt	60
tttttcttaa	atatcaccta	ttaggttgaa	aacctgaaat	tgcagctttc	tgtagaaatg	120
gcggaagaca	aactaacatt	tttaaagcgc	tctcatttag	ctctgatgag	tactacaccc	180
ctnatattct	tctgatacta	aaataatttt	cctagtgtag	tctaaacttt	tttaaaaaga	240
catgtaatcc	gcggagttag	taactcaaaa	cgagtgcac	tnggaagtat	cgcagccgtt	300
nctggatnaa	attcccagct	tgcctngctt	ctnagccggg	gggctggtna	aaaaacatct	360
gcagcccnng	ggnaaaaacc	ttcgatttgt	tcttacgtgt	ttacgttatt	ttatttcctt	420
nnagcaaggc	nggganttgg	ggactcgaaa	tggtagagtt	gggctgggga	tcgcccttgt	480
tacataaaag	ncgtccagaa	gagggacggt	tacaggcnng	ganctccaaa	ggtcagtcct	540
tgccatt						547

<210> 670

<211> 232

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(232)

<223> n = A,T,C or G

<400> 670

cgaactat	agactaccta	ggaaaattat	tttagtatca	gaagaatatc	aggggtgtag	60
tactcatcag	agctaaatga	gagcgcttta	aaaatgttag	tttgtcttcc	gccatttcta	120
cagaaagctg	caatttcagg	ttttcaacct	aataggtgat	atttaanaaa	aaaaaaaagc	180
aatcgcaaat	agccccactg	cttttacaaa	tcattttttc	cccaacacaa	tg	232

<210> 671

<211> 214

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(214)

<223> n = A,T,C or G

<400> 671

ctcccccttc	ntccttcgct	actnncatt	ttcnnaaatt	tntttcgcnt	atgnggaaaa	60
acaccacat	tnttcanctc	gcacagaaca	ngnnggggtg	tgtaaaatga	agggttccn	120
cnccttctct	tattnaanaa	cactnaaana	gggangggct	aaaaccgcg	ngatntctac	180
nctatcgcg	gcgcttttgg	ngttggctag	aaga			214

<210> 672

<211> 328

<212> DNA

<213> Homo sapien



```

<220>
<221> misc_feature
<222> (1)...(328)
<223> n = A,T,C or G

<400> 672
ngancagcgg ngtttaaacg ggcctctaga ctcgaggaga cncctgttgg atggtggatc      60
acanntcgnt actactatac aggacagagt atcggganct cttggntggt ggngcctgcc      120
aaccactgct nctgttaact gcgtatctga agggactcgg actggcttca gaagaactac      180
cggctcgaat gnaccatgga tgattcncnc tagttgaaaa aaaactcagg cacatgtatt      240
gccactgatg actagcgcca gactnctctc ggctctntaa cgagcccaca tgncngtggt      300
ncncccgtagc tgnctccaga agaggttc      328

<210> 673
<211> 223
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(223)
<223> n = A,T,C or G

<400> 673
ggggggcaaag ctgggctagcg tttaaactta agcttggtac cgagctcggga tcccnagac      60
attgtgcatg aaaatgcaaa ttgagtgtgg tctatantgc catcntcacc tncctgnengc      120
tcaaaaacaac ngctttctgc tgcaatgggt agggctcctn acncacgggtc gcnnacggag      180
gccnnccttat cctcntcggg nnggatccct ngaagcatnt tct      223

<210> 674
<211> 256
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(256)
<223> n = A,T,C or G

<400> 674
gnggggtcnt ngatgagcgc gcgtaatacn atcactntcn ggcnngntgg gtaccggggcc      60
ccccctcnaa gcggccgccc tttttttntt ttttttcatn acatgataan ntctttnttc      120
taaacagacc acaccactan agttcctttt ctttngtacg gaattgagtt aaagtagagn      180
atacaatgca gggcttcnnc tctatttcac attccaggnt ggttcngnat ggatcgggcc      240
tgctctccg atgggt      256

<210> 675
<211> 439
<212> DNA
<213> Homo sapien

```



<221> misc feature

<223> n = A, T, C or G

nnactagtc	agtgtggtg	aattccattg	tgttgggctt	gtatgggttt	ttttgtctag	60
ttntttggga	aatgttngtg	ttactatntt	ttggatatna	tatatgatat	gtatggccct	120
tctatgggct	cctcanacng	aactcaacca	ttttccacaa	aaccnattcc	tcctttccct	180
tcatgactga	gtggtgttg	tactatccng	gaaactggga	cattgtcctt	cacatctntc	240
ccttanctgc	ctngtccnat	tgatgtcttt	gagctntgan	atgtccttgt	taactntctc	300
ctnctctgt	actgccggca	naattaagca	ccatntgtca	caaaaagtat	tgcgttacct	360
tcacgnatct	gttngttnc	atncttgctg	cttctccngn	ggaaaatagg	ctnttctggc	420
aaccgaacng	aanaaatac					439

<211> 587

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

ngngggcctn	attaagcgcg	cgtaatacna	ctcactntgg	ggcgaattgg	gtaccgggnc	60
cccctcaagt	tnatntgccn	aacctctctt	ttggaataac	aaaagggtta	acacatatgt	120
cctcataggg	acgcgctttc	acacnttcct	gacngcttca	tanacntcat	tnctattttct	180
cctcagnaca	agttnaggcn	gaagggtgagg	canacnttat	aatttcaccatt	tcacaaatnc	240
ggaaagtgag	gctcaaaggg	nttaaaaaat	aacctgatac	aantcataga	gccgggtntct	300
ggaanaagca	ggagcaaagt	ccaggcatcc	tgatccaagc	tnggtccact	gccttccact	360
ctggagaggc	ttcatctccg	acaaaggaag	ggacntgagt	ggctgganaa	tctcatggga	420
taaagacctc	agnatttcat	gctcctggaa	atcccatggg	ttgaacaaca	ggtntttggc	480
ccgtggttct	ntccctttgn	ccatctttta	accttgggg	aaatgatggc	ntctntnagc	540
nttttttttn	aaaagqatnq	aaattqaatq	attattnqct	cattggg		587

<211> 444

<213> Homo sapien

<221> misc feature

<223> n = A,T,C or G

gtggggcatn	attaagcgcg	cgtaatacga	ctcactatag	ggcggaantg	ggtaccgggc	60
ccccctcgaa	gcggccgccc	tttttttttt	tttttactgt	ccaaactntc	tatngatnta	120
gttgaactgt	ncaacgattt	catgaaattc	tatacacana	gccttcaggt	ccagagagta	180



```
<210> 678
<211> 670
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(670)
<223> n = A,T,C or G
```

```
<210> 679
<211> 449
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(449)
<223> n = A,T,C or G
```

<210> 680



```
<220>
<221> misc_feature
<222> (1)...(670)
<223> n = A,T,C or G
```

```
<210> 681
<211> 494
<212> DNA
<213> Homo sapien
```

<400>	681
tcatggtgtc cacagtctga tgtgagcgca ttaaatttaa ggatctccgc ctttctcett	60
aaaactcagg acttggaat gancctagga agcgccccct ccctcccان ccanatccaa	120
gccccggacc gctgcnctc cagctgcgcc tagtgaaacc gccgaattcg aattcacact	180
cggngggccg gcgaaggtgt gcgcgcccg gggagcgccg gggcnagccc gagggactgc	240
aagccaanaa nggaggcatg ggtggcgggg ggcgccgtct gatccaggaa ggagcggagg	300
cgccgatcac acactcttna gacgcacctgc ccgcgcctgg ccagcgcgca gnctgcagga	360
cgcgcggagc aggaactcgc tggagtttgc caagccccan gnctctggaa agtntgtagc	420
tcccttttcg ancgnctctt ctggcccttt gggacgggtg tgtcattggg cgggggtctg	480
tataaggggg ggac	494

```
<220>
<221> misc feature
```



<223> n = A, T, C or G

tgatcattca	agcngtngnc	gnataacgat	tgctnagccc	aacctttcat	agggtcgttc	60
ctttgggaat	nggatgtcta	ttgaatggca	gggatagggg	cactcggcat	tcgcctctgg	120
tacagttttg	catatatatc	ctcatcgcga	gcgagcgtag	gggancgtta	agtttggggg	180
aatgccnccg	catgncctn	cggagctta	aacccccaac	aatnccatt	ttnaaaaaag	240
ntttnttant	taaaaaaaaa	aac				263

<213> Homo sapien

<223> n = A, T, C or G

cttggccgggc	atgcacagac	ntntttacgg	acachctact	ccaagngagc	ctgnanctgt	60
ctacgggtcaa	nctctaaggt	tngncantgc	cacanatggc	atagtcccgga	gggcggtnan	120
tctggantgc	tctctgcact	tgaacntaaa	gcgcntttca	aganaggngct	aatngcctgc	180
ctcttgacaa	cnaacaancc	cacaccnacc	tangaccctn	tangcaagga	ctggattctg	240
naaatgcaat	acaca					255

<213> Homo sapien

<223> n = A, T, C or G

acccctcatt	tcatgtgctt	ctattttcc	acatctttta	catgactaag	ggattaatga	60
aatcacctct	tcataatcat	gaccataatt	tcaccaaca	agtactcaag	tttggtgtta	120
gcactttatt	aatgcttacg	aattctctct	ctctccctct	ttctcttttc	cttagtcctt	180
gcacaataag	gatttttgaa	tgtataatat	catcttaggt	aagctttcat	atggttttgg	240
catatgaagc	ttatgactgt	cataagccat	accaagcctg	tggagtatgg	catgattttc	300
attacataat	ccaatgaaaa	tagacttatt	ttaaatccct	aactttgtag	ttttaatttg	360
tatttcacta	tcttgaaatt	aacagctagt	acttatccat	cacagcagtc	tcctactgac	420
atgaagcaag	ttgttgaaatg	cagtaganca	tgaatgaaag	catttaaatgt	tanacaaaaa	480
tgggtgatac	ccaagcattc	tgaattattt	gcatcaagga	atgggacatg	tacattagtg	540
gcatcatttc	taccaatatg	tgacttgaat	tgttttttta	aaaaaaggan	aatgantttc	600
tcaattttgct	ttaaaaaatt	ttnaaaaagt	tcaatggcat	gctgctttgt	ctggacttaa	660
ttttattaaca	attnttaanc	cttccttaag	gacanaattt	tgggtgttcag	gatcnccttg	720
aagggtctta	tttttnatan	nattccaaac	ccaaaagggtg	gtttaaaatg	ggnqqgttcc	780



```

ccccncnaaa atttggaacg gcttttttat atttaaaaaa nttncnttt gngtttgaaa      840
nctnaatacc aattaagggg gaattttacc tnccagtggg aaaaaaaaaac nctngccntt      900
naaaaaattc ccnggagnca at                                           922

```

<210> 685

<211> 531

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(531)

<223> n = A,T,C or G

<400> 685

```

tgaggctctg taaaactggt cctctgctag gcatacttca tattctctat attaaactca      60
tctttaattg gcatggaaga ttcatgttc caaatctcag atgaagatcc tatattggat      120
gcaattaagc ctggcagcgc cctcaaaaga cagtcttgct actgctagcc acagccagga      180
cacagtaaca gttccttcta gtgaccnag accataanaa atananatct aaagaattct      240
gactccaaag gcattagccc attcctggta ttgccaatta tgatagaaaa aattgccaaag      300
ctcctgggac atggaaaatac actcagtaca tttgagaact ggagaactan tttccaaaat      360
agtatgaaga catganggtg attgtagata tntgagtttg gagaanttga gggaaatcng      420
attacacatg tttactacaa gagatgttna taagtaaaga aggcctgata tacaatctaa      480
cagacnantg agataaatct taantcacia ctgacntccc ttttggggcg g                    531

```

<210> 686

<211> 336

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(336)

<223> n = A,T,C or G

<400> 686

```

ggngncctna tgagcgcgcg taatacgatc atatagggcg aattgggtac cgggcccccc      60
tcaagaacac tacaagctat gtccctctct canagagccc tgaantttta acatattgaa      120
agctctnatc ttgccaanaa actccactta acttcaaaac acaccctcca cacacatcat      180
gatcaactna gatcttactg aaccagaatc ctnaatggca tacttcagga acaggggtcc      240
anagaagcag ttctcaaant gcagctnaaa aagaaactga aaaccaatt catgcaanac      300
ctagggctta tttgagagca ttttccagtg cagatt                    336

```

<210> 687

<211> 271

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(271)



<400> 687

<210> 688

<211> 740

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (740)$ 

<223> n = A, T, C or G

<400> 688

tgatgaagcg	cgcgtnntac	nactcactat	nggggcgaan	tatgggtacc	gggnccccct	60
cgaagcgggcc	gccctttttt	tntttttttg	tgagagttta	aataaaatat	ttgagtttaa	120
tttaaagttt	gagtttaatt	aaaatatatg	gcataccca	agttgggctt	tgcanaaaga	180
acacttctca	ggaactgtta	gttggtgtac	caggaactca	gaagggtcct	gttattaaat	240
atatttgga	aatgcatgga	ttctctgaan	atcnctctgc	atgtgagcaa	cacttacatc	300
ncaaaccaaa	attggcattg	catacatnaa	ccaatatttc	ccaaacattt	ctggttatgg	360
cccaccccc	ttgtgtanta	cttattgctg	ttttttggaa	ccctggggaa	attacttaaa	420
atattcagct	ggaaattaca	ggcgttactt	ttaaggganc	aagaattaca	gtgactccca	480
aaattgcaag	tgttgattac	tatttaagaa	ccaagaatt	tgaagaaaat	tttgaaaagt	540
gaaaacngga	aatnttaaat	gacttctcaa	attttgaaaa	ctcnggnaaa	catctccact	600
ttggtncct	tcctttaaaa	attggctaaa	aattntttnt	tatncccacc	ccattggaan	660
tnccccccc	ctggaacaat	tggattcccc	tatttcctaa	aaaacggccn	ccccccccgg	720
gqnqaacncc	nacnttttqn					740

<210> 689

<211> 635

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (635)$ 

<223> n = A, T, C or G

<400> 689

actagtcacag	tgtggtggaa	ttccattgtg	ttgggattac	atatactttt	agcaattttt	60
aaagaagtgt	acaaagttga	gatgtttcct	gagctctcat	atatctgana	atgtcatttt	120
acatctcogt	cttcacctct	caaaacttct	ttcaattcct	tggctcttaa	tagtaatcaa	180
cacttgcaact	ctggagtcac	tgtaattcct	gctcctttac	agctacncct	gttattttcca	240
gctgaatat	tttagttatt	tcccaggggt	ccaaaaaaca	gcaataagta	ctacacaaaag	300
ggggtggggc	ataaccaqaa	atgtttggga	aatactggct	catgtatgca	atgccaaatc	360



tggtttgcna	ttgtantggt	gctcacatgc	agagtgaatc	ttcaaanaat	ccatgcattt	420
tccaaatata	tttaataaca	gggaaccttc	tganttcctg	gntacaccaa	ctaacagttc	480
ctgaaaaatg	ttctttctgc	aaaacccaac	ttggggatat	gccatatatt	ttaattaaac	540
tcaaacttta	aattaaactn	caattatttt	attttaaact	cctcaaaaaa	aaaaaaaaaa	600
agggggggcc	cttccaangg	ggggnccggt	tcccc			635

&lt;210&gt; 690

&lt;211&gt; 3923

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 690

acagaagaaa	tagcaagtgc	cgagaagctg	gcatcagaaa	aacagagggg	agattttgtgt	60
ggctgcagcc	gagggagacc	aggaagatct	gcatgggtgg	aaggacctga	tgatacagag	120
gaattacaac	acataactt	agtgtttcaa	tgaacaccaa	gataaataag	tgaagagcta	180
gtccgctgtg	agtctcctca	gtgacacagg	gctggatcac	catcgacggc	actttctgag	240
tactcagtgc	agcaaagaaa	gactacagac	atctcaatgg	caggggtgag	aaataagaaa	300
ggctgctgac	tttaccatct	gaggccacac	atctgctgaa	atggagataa	ttaacatcac	360
tagaaacagc	aagatgacaa	tataatgtct	aagtagtgac	atgtttttgc	acatttccag	420
cccttttaaa	tatccacaca	cacaggaagc	acaaaaggaa	gcacagagat	ccctgggaga	480
aatgcccggc	cgccatcttg	ggtcacgat	gagcctcgcc	ctgtgcctgg	tcccgttgt	540
gaggaagga	cattagaaaa	tgaattgatg	tgttccttaa	aggatgggca	ggaaaacaga	600
tcctgttgtg	gatattttatt	tgaacgggat	tacagatttg	aaatgaagtc	acaaagttag	660
cattaccaat	gagaggaaaa	cagacgagaa	aatcttgatg	gcttcacaag	acatgcaaca	720
aacaaaatgg	aatactgtga	tgacatgagg	cagccaagct	ggggaggaga	taaccacggg	780
gcagagggtc	aggattctgg	ccctgctgcc	taaactgtgc	gttcataacc	aaatcatttc	840
atatttctaa	ccctcaaaac	aaagctgttg	taatatctga	tctctacggg	tccttctggg	900
cccaacattc	tccatatatc	cagccacact	catttttaaat	atthagttcc	cagatctgta	960
ctgtgacctt	tctacactgt	agaataacat	tactcatttt	gttcaaagac	ccttcgtgtt	1020
gctgcctaata	atgtagctga	ctgtttttcc	taaggagtgt	tctggcccag	gggatctgtg	1080
aacaggctgg	gaagcatctc	aagatctttc	caggggtata	cttactagca	cacagcatga	1140
tcattacgga	gtgaattatc	taatcaacat	catcctcagt	gtctttgccc	atactgaaat	1200
tcatttccca	cttttggtgcc	cattctcaag	acctcaaaat	gtcattccat	taatatcaca	1260
ggattaactt	tttttttttaa	cctggaagaa	ttcaatgtta	catgcagcta	tgggaattta	1320
attacataatt	ttgtttttcca	gtgcaaagat	gactaagtc	tttatccctc	ccctttgttt	1380
gatttttttt	ccagtataaa	gttaaaatgc	ttagccttgt	actgaggctg	tatacagcac	1440
agcctctccc	catccctcca	gccttatctg	tcatcaccat	caaccctccc	cataccacct	1500
aaacaaaatc	taacttgtaa	ttccttgaac	atgtcaggac	atacattatt	ccttctgcct	1560
gagaagctct	tccttgctct	ttaaatctag	aatgatgtaa	agttttgaat	aagttgacta	1620
tcttacttca	tgcaaagaag	ggacacatat	gagattcatc	atcacatgag	acagcaaata	1680
ctaaaagtgt	aatttgatta	taagagttta	gataaatata	tgaaatgcaa	gagccacaga	1740
gggaatgttt	atggggcacg	tttgtaagcc	tgggatgtga	agcaaaggca	gggaacctca	1800
tagtatctta	tataatatac	ttcatttctc	tatctctatc	acaatatcca	acaagctttt	1860
cacagaattc	atgcagtgca	aatccccaaa	ggtaaccttt	atccatttca	tgggtgagtgc	1920
gctttagaat	tttggcaaat	catactgggc	acttatctca	actttgagat	gtgtttgtcc	1980
ttgtagttaa	ttgaaagaaa	tagggcactc	ttgtgagcca	ctttaggggt	cactcctggc	2040
aataaagaat	ttacaaagag	ctactcagga	ccagttgtta	agagctctgt	gtgtgtgtgt	2100
gtgtgtgtgt	gagtgtacat	gccaaagtgt	gcctctctct	cttgacccat	tatttcagac	2160
ttaaaacaag	catgttttca	aatggcacta	tgagctgcca	atgatgtatc	accaccatat	2220
ctcattattc	tccagtaaat	gtgataataa	tgtcatctgt	taacataaaa	aaagtttgac	2280
ttcacaaaag	cagctggaaa	tggacaacca	caatatgcat	aaatctaact	cctaccatca	2340



```

gctacacact gcttgacata tattgttaga agcacctcgc atttgtgggt tctcttaagc 2400
aaaataacttg cattaggtct cagctggggc tgtgcatcag gcggtttgag aaatattcaa 2460
ttctcagcag aagccagaat ttgaattccc tcatctttta ggaatcattt accaggtttg 2520
gagaggattc agacagctca ggtgctttca ctaatgtctc tgaacttctg tccctcttg 2580
tgttcatgga tagtccaata aataatgtta tctttgaact gatgctcata ggagagaata 2640
taagaactct gagtgatata aacattaggg attcaaagaa atattagatt taagctcaca 2700
ctgggtcaaaa ggaaccaaga tacaagaac tctgagctgt catcgtcccc atctctgtga 2760
gccacaacca acagcaggac ccaacgcagc tctgagatcc ttaaatacaag gaaaccagt 2820
tcatgagttg aattctccta ttatggatgc tagcttctgg ccatctctgg ctctcctctt 2880
gacacatatt agcttctagc ctttgcctcc acgactttta tcttttctcc aacacatcgc 2940
ttaccaatcc tctctctgct ctggtgcttt ggacttcccc acaagaattt caacgactct 3000
caagtctttt ctccatccc caccactaac ctgaatgcct agacccttat tttattaat 3060
ttccaataga tgctgcctat gggctatatt gcttttagatg aacattagat atttaaagct 3120
caagaggttc aaaatccaac tcattatctt ctctttcttt cacctccctg ctctctctcc 3180
tatattactg attgcactga acagcatggg cccaatgta gccatgcaaa tgagaaaccc 3240
agtggctcct tgtggtacat gcatgcaaga ctgctgaagc cagaaggatg actgattacg 3300
cctcatgggt ggagggggacc actcctgggc ctctgtgatt gtcaggagca agacctgaga 3360
tgctccctgc cttcagtgct ctctgcatct cccctttcta atgaagatcc atagaatttg 3420
ctacatttga gaattccaat taggaaactca catgttttat ctgccctatc aattttttaa 3480
acttgctgaa aattaagttt tttcaaaatc tgtccttgta aattactttt tcttacagt 3540
tcttggcata ctatatcaac tttgattctt tgttacaact tttcttactc ttttatcacc 3600
aaagtggctt ttattctctt tattattatt attttctttt actactatat tacgttggtta 3660
ttattttggt ctctatagta tcaatttatt tgatttagtt tcaatttatt tttattgctg 3720
acttttaaaa taagtgattc ggggggtggg agaacagggg agggagagca ttaggacaaa 3780
tacctaatac atgtgggact taaaacctag atgatgggtt gataggtgca gcaaacct 3840
atggcacacg tatacctgtg taacaaacct acacattctg cacatgtatc ccagaacgta 3900
aagtaaaatt taaaaaaaag tga 3923

```

<210> 691

<211> 882

<212> DNA

<213> Homo sapien

<220>

<221> misc\_feature

<222> (1)...(882)

<223> n = A,T,C or G

<400> 691

```

ttactcacta tagggctcga ggggccgctg aattctgctg cagtgaactg tgattatgtc 60
cctgcactcc agcctggatg acagaacacg atcatttctc taaagacaaa caaaaaacat 120
aaaataaaaac tagtataagg atagaagccc aggggttgatt taagtctgcg gaaatcataa 180
accataggtc agacttctca ttgatgaggt acttggtgggt tagaatacaa ttaggtatat 240
ttggtctaga aaccaggatg gaattagaga ataaaagact gagcaatagc atgttatagt 300
attagaaata ctatagaaat aggaaaagcc ctgattatga ctttggagtt ctgatccaac 360
atctgggatt atttagatat tttaaaggaa aacgatgact tttagctctc aggatgttag 420
tttctcaac cataaaatga agagcctcga aaagatttctg tttaccagat tatttctgaa 480
gtcaattcca gttctaaaat tccatcactg ngcactaagg caaattgaat tgaataaagt 540
attgggnatg cataaaatac tctattttta aaaangaata gtaattatcc attggnaaca 600
gacgcantca tccagncatc tctacacctg ncccatgncn tatgtagana tgtantctta 660
atcccttaac aaaccgattt tgcaaaggag cttanccttg gggacttgg tcanggcaac 720

```



tggtctactt	tnaagactca	tottcactta	ctgggcacca	aatncctacc	attgcatcaa	780
actgggggttc	ccatncaagg	caaaccctgn	gaaatcttta	atcccgaat	tggcgcccaa	840
ttttnggggg	tttccnaaaa	gaatcntccc	ccccgagggg	cc		882

&lt;210&gt; 692

&lt;211&gt; 235

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1) ... (235)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 692

ccgcactngt	aangnccgcc	agnngctgn	aantccgctn	agcncggatc	cactagtcca	60
ttgatggtaa	aagggtagct	tactggnatg	tccgnetgct	ccanganata	atacncagga	120
cttctcanag	cacttaatat	gttaatatata	aactncgnga	aaaaagatnt	tcnatgaanc	180
nttctcttta	ggaggtcagg	ngagaatagt	gttaatgnca	ttaagganag	aacga	235

&lt;210&gt; 693

&lt;211&gt; 383

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1) ... (383)

&lt;223&gt; n = A,T,C or G

&lt;400&gt; 693

nttatgtaag	aaatgtcata	tatcttttat	tttctttaaa	tcaaaataaa	tatgactttg	60
agcatcccat	cccattgccc	atcctatcag	aatggtagga	acatcaacac	aaataattag	120
taatgcaccg	catctacatt	cccattgctc	ctttacttct	tcagcattgc	ctaaaggcat	180
aatacacctt	taattaatta	attcagcctc	ctaattgcaca	ttaacaaagc	ccctgctaga	240
ctctgtccat	aatggnaaac	ctgnatgata	cttgatatta	acantttaag	gaatgctcat	300
ggattggtn	cagacttaaa	aaattgaggg	ggctgaanaa	aatctaangg	anaaatcatg	360
gaagcatttg	cacatattac	ata				383

&lt;210&gt; 694

&lt;211&gt; 204

&lt;212&gt; DNA

&lt;213&gt; Homo sapien

&lt;400&gt; 694

tctcttggt	ggcagcctg	aagggtggta	atgactcacc	aacgctacta	atccttcttc	60
actgtccctt	atctttttcc	ctcccaggct	cataactoga	ggttaaactc	tctttttatac	120
aagaaccctg	tctgatgaag	catcatttca	gaattttaag	tcaacttaca	aatgtggtat	180
tattcacatc	tgagtacaaa	ttta				204

&lt;210&gt; 695



```
<220>
<221> misc_feature
<222> (1)...(670)
<223> n = A,T,C or G
```

```
<210> 696
<211> 317
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1) ... (317)
<223> n = A,T,C or G
```

```
<210> 697
<211> 246
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(246)
<223> n = A,T,C or G
```



<400>	698						
agaaagtttc	cttttttttt	tttaatggtg	aaaagatata	cacatattta	gaattagcca		60
gctgggctca	gttttagatta	ttccaatttt	gttggcaaca	tccagagcat	cgtaatcagg		120
agccagtga	acatatctct	tcttctctcc	atcaggccaa	atcacggtgt	tgaccttggc		180
cacatcaatg	tcttagaact	tcttcacagc	ctgtttgatc	tggtgcttgt	tggtcttaac		240
atccacaatg	aacacaagt	tggtgtgtc	ttctatcttc	ttcgtggtga	ctcagtggtc		300
agcggaaact	tgatgatagc	gtagtgggtc	agcttgtatc	tcttgggagc	gctcttccaa		360
agatatcttg	gctgcctcgg	gagttgcagc	gtcttggggc	gccggaaggt	gggtgacgta		420
cggatcttct	ttttttgtgt	ggctgtggac	acctttcaac	actgtcttct	tggtctttaa		480
atccttcgct	ttgggttcgg	ctataggagg	ggcaggagct	tcttcttcca	cttctggcgc		540
catcttgtga	aaagggaaag	tttcccttct	aataccattt	tcaactctcc	cgaattttgt		600
ggatcgtttc	ttgggtatcta	cccagattt	caggagtgtt	ggctggatct	tagggattgt		660
gaagtcttca	tttccctgtg	gtgagatctg	aggcatgatt	ttaaacagt	tgagggagg		720
agatctccag	gcactttaat	agaatggaga	agcaggatgg	gatttgagag	gaaatctgat		780
tttgaaaaaa	ggagaactag	agttgagttc	gtaattaaact	agcaccttaa	aggtcattca		840
gcatgcccat	ctgcacagt	gggttaatca	ccctacagaa	caaaaacaaa	aaggcaatgg		900
agaggaagct	gtaaagcact	gtacatgttt	aactcattgt	tatgtaagct	agccgaaggc		960
ttcacagact	tgaattcatc	tcccaagttc	tcttctgtga	ctggaaactc	tgccttaggt		1020
tgcttaaaa	ttgagaaaac	gaatattgct	tccctgcct	gccttcttga	gtacacttgc		1080
ctacacaaa	atgcacatcc	ttgtttgtgt	gtgtgtgtcc	atttgetgtg	acattcttgt		1140
gaaagtcaa	gtttcccagc	tggtgacata	cacaagtttg	tttggtgcaa	cctgtcagat		1200
gcatccctta	gacaggccct	ttgatactct	gggaaagaca	ttggacttac	agtcggaacg		1260
aaaagaaaga	aatgtgatat	gtatagcgtg	cagtgaagttg	gagttttacc	tgtattgttt		1320
taattttca	aagcctgagg	actagccaca	aatgtaccca	gtttacaaat	gaggaaacag		1380
gtgcaaaaag	gttgttacct	gtcaaaggtc	gtatgtggca	gagccaagat	ttgagcccag		1440
ttatgtctga	tgaacttagc	ctatgctctt	taaacttctg	aatgctgacc	attgaggata		1500
tctaaactta	gatcaattgc	attttccctc	caagactatt	tacttatcaa	tacaataata		1560
ccacctttac	caatctattg	ttttgatacg	agactcaaat	atgccagata	tatgtaaaag		1620
caacctacaa	gctctctaatt	catgctcacc	taaaagattc	cgggatcta	ataggctcaa		1680
agaaacttct	tctagaaaata	taaaagagaa	aattggatta	tgcaaaaatt	cattattaat		1740
ttttttcatc	catcctttta	ttcagcaaac	atttatctgt	tgttgacttt	atgcagtatg		1800
gcctttttaag	gattggggga	cagggtgaaga	acggggtgcc	agaatgcac	ctcctactaa		1860
tgaggtcagt	acacatttgc	atttttaaat	gccctgtcca	gctgggcacg	gtggatcatg		1920
cctgtaatct	caacattgga	aggccaaggc	aggaggattg	cttcagccca	ggagttcaag		1980
accagcctgg	gcaacataga	aagaccccat	ctctcaatca	atcaatcaat	gccctgtctt		2040
tgaaaataaaa	actctttaag	aaaggtttta	tgggcagggt	gtggtagctc	atgcctataa		2100
tacagcactt	tgaggaggctg	aggcaggagg	atcacttttag	cccagaagtt	caagaccagc		2160
ctgggcaaca	agtgcacct	catctcaatt	ttttaataaaa	atgaatacat	acataaggaa		2220
agataaaaaag	aaaagtttaa	tgaaagaata	cagtataaaa	caaactctctt	qqacctaaaa		2280







```
<210> 700
<211> 2841
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(2841)
<223> n = A,T,C or G
```

<400>	700						
gcagagcaca	gcatagctgc	tttaccaaat	catggccaga	ctgcttctgt	aagcaggccc		60
ctgatcctgt	tccacctcac	tggacaggac	ctcccaactg	gggcctccag	ctacccccac		120
cagcatccct	tggccaatgg	aaatttgaaa	tgttcctggg	acagagctcc	tggagagagg		180
ggcaggccac	cacctttgct	gtttgggtga	ctagccgttc	tggcctgcag	gctttggaga		240
gcccaagctg	acaaggggta	gaagaggtgc	ctcagcacag	cacagccacg	ctacgaaaac		300
atggccagac	tcttgtttaa	gtcagtcccc	gaacacattt	ctagtcagtg	ggtgaagtct		360
ttcaaccagg	gtctctggct	accttgactg	ctgttctctg	gccgacagag	gtctcaggcc		420
tccctgagtc	agagctcccc	gggggaggac	cagattgtca	tctttgctgt	ttgggtgacc		480
cagccatttc	agccttaggg	cttcagagtg	tctgaggtag	ccaggggctg	aagtgaaccc		540
ccagcacagc	acagctgctg	tataaaaaacg	tggccagact	ttttcttttaa	gcaagtcctt		600
gttcttattc	ctcctgacta	ggtaagactt	ctcaacttgc	ctccagccac	atcttattgg		660
tgtgttcaga	ttggcaacag	gtttgtacct	cagtgggtaca	gagctcccag	aggaaggggt		720
aggctatcat	cttccctgga	aaatacagat	caattagggg	cttgagggga	ccccagcat		780
tccacagcag	cccttcagaa	aagtggccag	actctgtact	tgatggggcag	atcctcctgg		840
cctgtgtctc	tagccagccc	accactggag	ctatcaagcc	agtagcaact	cagcagttcc		900
ttggacagag	cttccaggag	caaatagaat	cctttctgcc	actgcctttg	cagtgaactg		960
cccttgctat	cctcagaaga	tatatcacgg	gagcaaagac	cctaagtgcc	atatcaacac		1020
ctccaataag	ctgcagttga	cccaaagaac	aagccaatcc	atctcccaca	ggttccacac		1080
acactccact	actcatcacc	aqacagggaa	ccctggcttg	ggcccacagc	acaqaccctc		1140



```
<210> 701
<211> 3228
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(3228)
<223> n = A,T,C or G
```

```
<400> 701
tccgccccat tgacgcaaat ggcggtaggc gtgtacggtg ggaggtctat ataagcagag
60aactagagaa ccactgctt actggcttat cgaaattaat acgactcact      120
atagggagac ccaagctggc tagcgtttaa acttaagctt ggtaccgagc tcggatccac      180
tagtccagtg tgggtgaatt ccattgtgtt gggcaggaaa caagcaaagt ggtggagcag      240
caagtcaggt gatgtggagc ccagaggtca gggatggctg tctctctagg gtccacttgc      300
ccttgtgaga cactttatcc cagcacttta ggaatactga ggtcatacca gccacatctt      360
atatgcaaga ttgccagca gagatcaggt ccgagagttc cctttttaaa aaaaggagac      420
ttgcttaata aaagaagtct agccacgttt gtgtagagcg gctgtgctgt gctggggggt      480
cacttttgag agagttctcc tctgagacct gatctctgga ggctgggcaa tcttgcactt      540
gagatggggc tgggtctgat tcagcactcc ttagtctgct cgcctctccc atggccccaq      600
```







gtgcacgcgc	tggctcgcgg	tgacagccgc	gcgcctcggc	caggatctga	gtgatgagac	60
gtgtccccac	tgaggtgccc	cacagcagca	ggtgttgagc	atgggctgag	aagctggacc	120
ggcaccaaag	ggctggcaga	aatgggcgcc	tggctgattc	ctaggcagtt	ggcggcagca	180
aggaggagag	gccgcagctt	ctggagcaga	gccgagacga	agcagttctg	gagtgcctga	240
acggccccct	gagccctacc	cgcttgcccc	actatggctc	agaggctgtg	ggtgagccgc	300
ctgctgcggc	accggaagc	ccagctcttg	ctgggtcaacc	tgctaacctt	tggcctggag	360
gtgtgttttg	ccgcaggcat	cacctatgtg	ccgcctctgc	tgctggaagt	gggggtagag	420
gagaagttca	tgaccatggt	gctgggtgag	tactacatc	ctccttcctt	cctgttccag	480
atacatgcca	cctggcatgt	gggacaggag	tacctctgcc	ctgggagctg	cttggaggga	540
gaggtgggtct	gctgggaagg	cattgctggg	caggagggtg	accctgggct	gagggggcac	600
accaagagaa	agaagagaat	accaaggaca	tacccagtc	acctctggat	ccctggctct	660
gcacagagcc	tggctcatag	gagacactgg	agaaatgtct	ctaacctttg	gctagccctt	720
ttataattta	tagcgattat	ctcatttaat	gcttacaacc	accatttgag	gtgatccatt	780
ttacagagaa	ggaagcagag	gcttttaaga	ggttaggtaa	gtcttagcca	aagccaaata	840
gcagctgaac	agtagagctg	ggactccatc	aagggtctcc	agccggagct	tgctcctacc	900
cctaggacaa	gggggtggact	cctgactctg	cagataaatt	ctacaaaagc	cacagaaggc	960
aagtagtaac	catttgtgtga	caaccctca	ccccaggaa	gaggggcccc	tgtgaggatt	1020
gcaggctctg	gagtcacact	gcttgttgaa	acgtgcctc	ttaccctccc	taggtctgcg	1080
cctttgaata	agtatcactt	cttagttgct	ccatgcctca	gtttgtccat	ctgaaaatgg	1140
gggcatctgt	aatgcctgtg	ttatgaggag	taaattacag	catccctgtg	aagacgtagc	1200
acagtgtcga	gtacggaatg	ttatttccat	ccttctcacg	gagcttggtt	ccccttcccc	1260
ttgcoccttta	cttgtcccag	ccattgactc	atactacttc	ccttcttgca	ggcattgggtc	1320
cagtgtctggg	cctgggtctgt	gtcccgtctc	taggtctcagc	cagtgaccac	tggcgtggac	1380
gctatggccg	ccgccggccc	ttcatctggg	cactgtcctt	gggcatcctg	ctgagcctct	1440
ttctcatccc	aagggccggc	tggctagcag	ggtgctgtg	ccgggatccc	aggcccttg	1500
agctggcact	gctcatcctg	ggcgtggggc	tgtgtggactt	ctgtggccag	gtgtgcttca	1560
ctccactgga	ggccctgtct	tctgacctct	tccgggaccc	ggaccactgt	cgccaggcct	1620
actctgtcta	tgcttcatg	atcagtcttg	ggggtgcct	gggctacctc	ctgcctgcca	1680
ttgactggga	caccagtgc	ctggccccct	acctgggcac	ccaggaggag	tgctcttttg	1740
gcctgtcac	cctcatcttc	ctcacctgcg	tagcagccac	actgctggtg	gctgaggagg	1800
cagcgtggg	ccccaccgag	ccagcagaag	ggtgttcggc	ccctccttg	tcgccccact	1860
gctgtccatg	ccgggcccgc	ttggctttcc	ggaacctggg	cgccctgctt	ccccggctgc	1920
accagctgtg	ctgccgcag	ccccgcaccc	tgcgcggct	cttcgtggct	gagctgtgca	1980
gctggatggc	actcatgacc	ttcacgctgt	tttacacgga	tttcgtgggc	gaggggctgt	2040
accagggcgt	gcccagagct	gagccgggca	ccgaggcccg	gagacactat	gatgaaggta	2100
aggccttggc	agccagcaga	ggctgggtgtg	ggagccgcct	accagagacg	acactcgggg	2160
ctgtgtcttg	gctgggtgct	ctccatcctg	gccccgactt	ctctgtcagg	aaagtgggga	2220
tggaccccat	ctgcatacac	ggcttctcat	gggtgtggaa	catctctgct	tgcggtttca	2280
ggaaggcctc	tggctgtctc	aggagtctga	tcagagtctg	tgccccagtt	tgacagaagg	2340
aaaggcggag	cttattcaaa	gtctagaggg	agtggaggag	ttaaggctgg	atttcagatc	2400
tgctctgggtc	cagccgcagt	gtgcctctg	ctccccaac	gactttccaa	ataatctcac	2460
cagcgccttc	cagctcaggc	gtcctagaag	cgtcttgaag	cctatggcca	gctgtctttg	2520
tgttccctct	caccgcctg	tcctcacagc	tgagactccc	aggaaacctt	cagactacct	2580
tcctctgcct	tcagcaaggg	gcgttgccca	cattctctga	gggtcagtgg	aagaacctag	2640
actcccattg	ctagaggtag	aaagggggaag	ggtgctgggg	agcagggtg	gtccacagca	2700
ggtctcgtgc	agcaggtacc	tgtygttccg	ccttctcatc	tccttgagac	tgctccgacc	2760
cttccctccc	aggctctgtc	tgatggcccc	tctccctctg	caggcgttcg	gatgggcagc	2820
ctggggctgt	tcctgcagtg	cgcatctctc	ctgggtcttct	ctctggctcat	ggaccggctg	2880
gtgcagcgat	tcggcactcg	agcagtctat	ttggccagtg	tggcagcttt	ccctgtggct	2940
gccgggtgcca	catgcctgtc	ccacagtgtg	gccgtgggtga	cagcttcagc	cgccctcacc	3000



```

gggttcacct tctcagccct gcagatcctg ccctacacac tggcctccct ctaccaccgg 3060
gagaagcagg tgttcctgcc caaataccga ggggacactg gaggtgctag cagtgaggac 3120
agcctgatga ccagcttcct gccaggccct aagcctggag ctcccttccc taatggacac 3180
gtgggtgytg gaggcagtg cctgctccca cctccaccog cgtctctcgg ggctctgcc 3240
tgtgatgtct ccgtacgtgt ggtggtgggt gagccaccog aggccagggt ggttccgggc 3300
cggggcatct gcttgacct cgccatcctg gatagtgcct tcctgctgtc ccagggtggcc 3360
ccatccctgt ttatgggctc cattgtccag ctccagccagt ctgtcactgc ctatatggtg 3420
tctgccgcag gctgggtct ggtcgccatt tactttgcta cacaggtagt atttgacaag 3480
agcgacttgg ccaaatactc agcgtagaaa acttcacgca cattgggggtg gagggcctgc 3540
ctcactgggt ccagctccc tgctcctgtt agcccatgg ggctgcggg ctggccgcca 3600
gtttctgttg ctgcaaaagt aatgtggctc tctgctgcca ccctgtgctg ctgagggtgcg 3660
tagctgcaca gctgggggct ggggcgtccc tctcctctct cccagctctc tagggctgcc 3720
tgactggagg ccttccaagg gggtttcagt ctggacttat acagggaggc cagaagggtc 3780
ccatgcactg gaatgcgggg actctgcagg tggattaccc aggtcagggt ttaacagcta 3840
gcctcctagt tgagacacac ctagagaagg gtttttggga gctgaataaa ctcagtcacc 3900
tggtttccca tctctaagcc ccttaacctg cagcttcgtt taatgtagct cttgcatggg 3960
agtttctagg atgaaacact ccacatggg atttgaacat atgaaagtta tttgtagggg 4020
aagagtccct aggggcaaca cacaagaacc aggtcccctc agccacagc actgtctttt 4080
tgctgatcca cccctctt accttttatc aggatgtggc ctggttggtc ttctgttgcc 4140
atcacagaga cacaggcatt taaatattta acttatttat ttaacaaagt agaagggaat 4200
ccattgctag ctttctgtg ttggtgtcta atatttgggt aggggtgggg atccccaaca 4260
atcagggtccc ctgagatagc tggtcattgg gctgatcatt gccagaatct tcttctcctg 4320
gggtctggcc ccccaaatg cctaaccag gaccttgga attctactca tcccaaatga 4380
taattccaaa tgctgttacc caaggttagg gtgttgagg aaggtagagg gtggggcttc 4440
aggtctcaac ggcttccta accacccctc ttctcttggc ccagcctggg tccccccact 4500
tccactcccc tctactctct ctaggactgg gctgatgaag gcaactgcca aaatttcccc 4560
tcccccaac tttccctac ccccaacttt cccaccagc tccacaacc tgtttgaggc 4620
tactgcagga ccagaagcac aaagtgcggt ttcccaagcc tttgtccatc tcagcccca 4680
gagtatatct gtgcttgggg aatctcacac agaaactcag gagcaccccc tgctgagct 4740
aaggagggtc ttatctctca gggggggttt aagtgcggt tgcaataatg tcgtcttatt 4800
tatttagcgg ggtgaatatt ttatactgta agtgagcaat cagagtataa tgtttatggt 4860
gacaaaatta aaggctttct tatatgttta aaaa 4894

```

<210> 703

<211> 2904

<212> DNA

<213> Homo sapiens

<400> 703

```

gtctatgcct tcatgatcag tcttgggggc tgctgggct acctcctgcc tgccattgac 60
tgggacacca gtgccctggc cccctacctg ggcacccagg aggagtgcct ctttggcctg 120
ctcaccctca tcttcctcac ctgcgtagca gccacactgc tgggtggctga ggaggcagcg 180
ctgggcccc cagagccagc agaagggctg tcggccccct ccttgtcgc ccactgctgt 240
ccatgcgggg ccgcttggc tttccggaac ctgggcgccc tgcttccccg gctgcaccag 300
ctgtgctgcc gcatgcccc caccctgcgc cggctcttcg tggctgagct gtgcagctgg 360
atggcactca tgaccttcac gctgttttac acggatttcg tgggcgagg gctgtaccag 420
ggcgtgcccc gagctgagcc gggcaccgag gcccgagac actatgatga aggaaggcct 480
ctggctgctc taggagtctg atcagagtgc ttgccccagt ttgacagaag gaaaggcgga 540
gcttattcaa agtctagagg gagtggagga gttaaggctg gatttcagat ctgcctgggt 600
ccagccgcag tgtgccctct gctcccccac cgactttcca aataatctca ccagcgcctt 660

```



```
<210> 704
<211> 4034
<212> DNA
<213> Homo sapiens
```

<400> 704						
aaccagcctg	cacgcgctgg	ctccgggtga	cagccgcgcg	cctcggccag	gatctgagtg	60
atgagacgtg	tccccactga	ggtgccccac	agcagcaggt	gttgagcatg	ggctgagaag	120
ctggaccggc	accaaagggc	tggcagaaat	gggcgcctgg	ctgattccta	ggcagttggc	180
ggcagcaagg	aggagaggcc	gcagcttctg	gagcagagcc	gagacgaagc	agttctggag	240
tgcttgaacg	qccccctgaq	ccctaccgcq	ctggcccaact	atggtccaqa	qqctgtqqqt	300







```

ttggtgtcta atatttgggt aggggtggggg atcccccaaca atcaggtccc ctgagatagc 3420
tggtcattgg gctgatcatt gccagaatct tcttctcctg ggggtctggcc ccccaaatg 3480
cctaaccag gaccttggaa attctactca tcccaaatga taattccaaa tgctgttacc 3540
caaggttagg gtgttgaagg aaggtagagg gtgggggttc aggtctcaac ggcttcccta 3600
accacccctc ttctcttggc ccagcctggg tccccccact tccactcccc tctactctct 3660
ctaggactgg gctgatgaag gcactgcccc aaatttcccc taccaccaac tttccctac 3720
ccccacttt cccaccagc tccacaaccc tgtttggagc tactgcagga ccagaagcac 3780
aaagtgcggt ttcccaagcc tttgtccatc tcagcccccga gagtatatct gtgcttgggg 3840
aatctcacac agaaactcag gagcaccccc tgccctgagct aagggagggtc ttatctctca 3900
gggggggttt aagtgccgtt tgcaataatg tcgtcttatt tatttagcgg ggtgaatatt 3960
ttatactgta agtgagcaat cagagtataa tgtttatggg gacaaaatta aaggctttct 4020
tatatgttta aaaa 4034

```

<210> 705

<211> 6976

<212> DNA

<213> Homo sapiens

<400> 705

```

gaagctggac cggcaccaaaa gggctggcag aaatggggcg ctggctgatt cctaggcagt 60
tggcggcagc aaggaggaga ggccgcagct tctggagcag agccgagacg aagcagttct 120
ggagtgcctg aacggccccc tgagccctac ccgcctggcc cactatgggtc cagaggctgt 180
gggtgagccg cctgctgcgg caccggaaag ccagctctt gctgggtcaac ctgctaacct 240
ttggcctgga ggtgtgtttg gccgcaggca tcacctatgt gccgcctctg ctgctggaag 300
tgggggtaga ggagaagttc atgacctagg tgctgggtga gtcactacat cctccttctc 360
tcctgttcca gatacatgcc acctggcatg tgggacagga gtacctctgc cctgggagct 420
gcttggaggg agagggtggtc tgctgggaag gcattgctgg gcaggagggt gaccctgggc 480
tgagggggca caccaagaga aagaagagaa taccaaggac ataccacagt cacctctgga 540
tcctgggtcc tgcacagagc ctggctcata ggagacactg gagaaatgct cctaaccttt 600
ggctagccct tttataattt atagcgatta tctcatttaa tgcttacaac caccatttga 660
ggtgatccat tttacagaga aggaagcaga ggcttttaag aggttaggta agtcttagcc 720
aaagccaaat agcagctgaa cagtagagct gggactccat caaggtctcc cagccggagc 780
ttgctcctac ccctaggaca aggggtggac tcctgactct gcagataaat tctacaaaag 840
ccacagaagg caagtagtaa ccattgtgtg acaaccctc accccagga agagggggcc 900
ctgtgaggat tgcaggctct ggagtcacac tgcttgttga aacgctgcct cttacctcc 960
ctaggctctgc gcctttgaat aagtatcact tmttagttgc tccatgcctc agtttgtcca 1020
tctgaaaatg ggggcatctg taatgcctgt gttatgagga gtaaattaca gcatccctgt 1080
gaagacgtag cacagtgtcg agtacggaat gttatttcca tccttctcac ggagcttggg 1140
tcccttccc cttgcccttt acttgtccca gccattgact catactactt cccttcttgc 1200
aggcattggg ccagtgtctg gcctgggtctg tgtcccgtc ctaggctcag ccagtacca 1260
ctggcgtgga cgctatggcc gccgcggcc cttcatctgg gcactgtcct tgggcatcct 1320
gctgagcctc tttctcatcc caagggccgg ctggctagca gggctgctgt gcccgatcc 1380
caggccctg gagctggcac tgcctatcct gggcgtgggg ctgctggact tctgtggcca 1440
ggtgtgtctc actccactgg aggcctgtct ctctgacctc ttccgggacc cggaccactg 1500
tcgccaggcc tactctgtct atgccttcat gatcagctct gggggctgcc tgggctacct 1560
cctgcctgcc attgactggg acaccagtgc cctggccccc tacctgggca cccaggagga 1620
gtgcctcttt ggctgtctca cctcatctt cctcacctgc gtagcagcca cactgtgtgt 1680
ggctgaggag gcagcgtctg gccccaccga gccagcagaa gggctgtcgg cccctcctt 1740
gtcgccccc tgctgtccat gccggggccg cttggctttc cggaacctgg gcgcctgtct 1800
tcccggctg caccagctgt gctgcgcgat gcccgcacc ctgcgcggc tcttcgtggc 1860

```



ttagctgtgc	agctggatgg	cactcatgac	cttcacgctg	ttttacacgg	atttcgtggg	1920
cgaggggctg	taccagggcg	tgcccagagc	tgagccgggc	accgaggccc	ggagacacta	1980
tgatgaaggt	aaggccttgg	cagccagcag	aggctggtgt	gggagccgcc	caccagagac	2040
gacactcggg	gctgtgtctg	ggctggtgcc	tctccatcct	ggccccgact	tctctgtcag	2100
gaaagtgggg	atggacccca	tctgcataca	cggcttctca	tgggtgtgga	acatctctgc	2160
ttgcggtttc	aggaaggcct	ctggctgtct	taggagtctg	atcagagtcg	ttgccccagt	2220
ttgacagaag	gaaaggcgga	gcttattcaa	agtctagagg	gagtgaggga	gttaaggctg	2280
gatttcagat	ctgcctggtt	ccagccgcag	tgtgccctct	gctcccccaa	cgactttcca	2340
aataatctca	ccagcgccct	ccagctcagg	cgctcctagaa	gcgtcttgaa	gcctatggcc	2400
agctgtcttt	gtgttccctc	tcacccgcct	gtcctcacag	ctgagactcc	caggaaacct	2460
tcagactacc	ttcctctgcc	ttcagcaagg	ggcgttgccc	acattctctg	agggtcagtg	2520
gaagaaccta	gactccatt	gctagaggta	gaaaggggaa	gggtgctggg	gagcagggct	2580
ggtccacagc	aggtctcgtg	cagcaggta	ctgtgggttc	gccttctcat	ctccctgaga	2640
ctgctccgac	ccttccctcc	caggctctgt	ctgatggccc	ctctccctct	gcaggcgttc	2700
ggatgggcag	cctggggctg	ttcctgcagt	gcgccatctc	cctgggtcttc	tctctggtca	2760
tggaccggct	ggtgcagcga	ttcggcactc	gagcagtcct	tttggccagt	gtggcagctt	2820
tcctctgtggc	tgccggtgcc	acatgcctgt	cccacagtgt	ggcctgtgtg	acagcttcag	2880
ccgccctcac	cgggttcacc	ttctcagccc	tgcagatcct	gccctacaca	ctggccctccc	2940
tctaccaccg	ggagaagcag	gtactcattg	gccagtgggt	ggagtcaggg	tgggaggggt	3000
ggtctgggtt	tttgggaggc	caactagctc	agaacctggt	atctggcaag	caactttgga	3060
gaatgcttct	ttgaatcaga	gaagaagctt	atcctagccc	cagggccaga	ggcttgggct	3120
gcagaacagt	gtagattaga	ttctgggaat	gacttcctgg	ggtcaggact	gtgtagcact	3180
tgaatggatg	attgcaggaa	atgcaaaata	cgatagtggg	aatcccgaag	ggtcaggcca	3240
gcaggagccc	taggcttcta	ggctggttgt	tctatggaga	ggcagggcgc	tgaatcagat	3300
gacccctggg	ccattcagcc	tcagcagacg	ggagtgggaa	tggctccagcc	ttagcaacac	3360
ctttcttcag	ggagcagcaa	cctgacttag	cctgtatcct	actctgggtct	ctgagatggg	3420
gcaggctcct	tctaccccc	tttctttctg	gcttattttt	cttttctgtc	taattccctt	3480
ttcttttctc	gcacccctcc	tttgctcctc	tccctttctc	cttccccctc	cccttccctt	3540
gtggcagata	tctgagcttg	acacctgacc	cactcaactg	ggcaactgtg	aagttgtggg	3600
gacctccttc	ttggttggcc	ctacactaac	cagccctccc	aggggcccct	ttccttggga	3660
agccacctaa	cccaggtagt	gtggtcatcc	ttgtccctcc	cactgacctc	actgagctac	3720
aaacctgggt	gctggactct	gccttgaggg	gcatgaagtt	ggggtgtccc	aaggaggagg	3780
gagatgcagg	actgctctca	tagagctctc	agactgtagg	gaagacctgc	ccctgcgtct	3840
cgtagcactt	gaggagagga	gtaggtaagt	tcgtagctga	gaggctggtt	aactgagtag	3900
gtagctgcag	gggtgagagg	tatggagggg	aggggctaag	gttttgggtg	ggggagcctg	3960
gtccctgaga	cccctgttag	cccactgata	accttcttca	gccttcactc	ttctgcttgc	4020
ctgggctggg	ggcagggggc	tggcatcagc	ggccaggcct	gagtatgtgc	tgtcgtgcca	4080
gggaacgttc	tggggctagc	catcttctcc	agatggagga	gcatgtctgt	cctcggacca	4140
ctccagactc	caacctcagc	ggacattcct	ggggtggcag	gcaggaggga	gaagtccctg	4200
gaggccccct	cctaacagca	gctgatggca	gacttggcac	tgcacgctgt	ctgcctgttc	4260
ctttgcccac	ttgttgagct	gcatggtgag	ccgtgggctt	ccctggtgtc	aggtttgagc	4320
tctgccatgg	ctcccaacct	gcaaatgcag	ccaactcaac	tcttctggca	tggggacaat	4380
gttgataaag	acctggcctt	gtccttaa	aggaggctct	gggccatcaa	gggcaggggt	4440
tggggggatg	gtggtcgacc	agtcactctg	atctaagtca	gacagcagga	aggaagttag	4500
aagccttcaa	cattagcaca	gctggggctg	ggggagggtg	gaagagggac	attcctcctg	4560
cttggggctc	actggattct	ccctgcccc	aggctgggga	caaggagact	catggcaggg	4620
cagctaccct	agtggcatct	gggaccccag	agaggcagag	cttctctgca	ccgggcaatg	4680
aggatttcca	gatgtcggag	tggagggcag	gcaggaagga	aggttaggag	agcctgcgtg	4740
gggtttgggc	catcaggggc	cctgccttgg	cttttgttcc	tctgttctgt	gcatctctta	4800
ccaccgtctt	cattccccct	gtgtcttttc	cttaccttgg	agctctgttc	tctctgatct	4860
gtgatattga	gtttgtctgc	ctcttacctg	ttctaagagg	ctaaaggaqa	cctagacttc	4920



Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys  
35 40 45







145

150

<210> 708

<211> 371

&lt;212&gt; PRT

<213> Homo sapiens

<400> 708

Met Leu Phe Pro Ser Phe Ser Arg Ser Leu Val Pro Leu Pro Leu Ala  
5 10 15

Leu Tyr Leu Ser Gln Pro Leu Thr His Thr Thr Ser Leu Leu Ala Gly  
20 25 30

Ile Gly Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala  
35 40 45

Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp  
50 55 60

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala  
65 70 75 80

Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu  
85 90 95

Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp Phe Cys Gly Gln Val  
100 105 110

Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro  
115 120 125

Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu  
130 135 140

Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser  
145 150 155 160

Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu  
165 170 175

Leu Thr Leu Ile Phe Leu Thr Cys Val Ala Ala Thr Leu Leu Val Ala  
180 185 190

Glu Glu Ala Ala Leu Gly Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala  
195 200 205

Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe  
210 215 220



Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His Gln Leu Cys Cys Arg  
 225 230 235 240

Met Pro Arg Thr Leu Arg Arg Leu Phe Val Ala Glu Leu Cys Ser Trp  
 245 250 255

Met Ala Leu Met Thr Phe Thr Leu Phe Tyr Thr Asp Phe Val Gly Glu  
 260 265 270

Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr Glu Ala Arg  
 275 280 285

Arg His Tyr Asp Glu Gly Lys Ala Leu Ala Ala Ser Arg Gly Trp Cys  
 290 295 300

Gly Ser Arg Pro Pro Glu Thr Thr Leu Gly Ala Val Ser Gly Leu Val  
 305 310 315 320

Pro Leu His Pro Gly Pro Asp Phe Ser Val Arg Lys Val Gly Met Asp  
 325 330 335

Pro Ile Cys Ile His Gly Phe Ser Trp Val Trp Asn Ile Ser Ala Cys  
 340 345 350

Gly Phe Arg Lys Ala Ser Gly Cys Ser Arg Ser Leu Ile Arg Val Val  
 355 360 365

Ala Pro Val  
 370

<210> 709

<211> 141

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(141)

<223> n=A,T,C or G

<400> 709

tacggcgtgg tgcggagggc ggtaccccccac aaataacacn nacaccccat cctatctgtg 60  
 tccacanata aantgactca ttctctctct cgcatanccc actntcccct ngcgataccg 120  
 taacnaancc cttccccctt t 141

<210> 710

<211> 196

<212> DNA

<213> Homo sapiens



<223> n=A,T,C or G

```

cnaatccttcn cntacaccca tgangtccat gtgcgacgtc cacctcccct caaaacttgg 60
gtccncatcc acccgtcact ctcccntaa ncnataaccc cttttngcga atagacccca 120
ccttancaat nggtttttcn ttttttgtec ctnggnccgn gcgattcaan aaattgaagg 180
cccanaaaaa cccctt                                     196

```

<213> Homo sapiens

<223> n=A, T, C or G

ntacntcnc	ccnaatgaaa	ttcgaanctc	ggttaccccg	gggnattccg	attaggngcg	60
tantctcgga	tgtgcagtea	caagtctttt	gctaattnc	ataattntcn	ctaccctttc	120
ttcnacaata	ctgtatctct	antntttctn	tcncctctct	cccannttac	taaccac	177

<213> Homo sapiens

<223> n=A,T,C or G

```

aaacgnacca  nngccaacga  tangtggttg  ngttggttgc  ggttgttcct  cttatntgca  60
ctggttgtcc  gtgtcgcacg  ganggccacg  tccctctgnc  ntgagtanca  catagcatcc  120
acgttttagtc  gactntnccg  ggcggccgct  ctaccctnt  atngattcct  attaaaantc  180
ggatc                                             185

```

<213> Homo sapiens

<223> n=A,T,C or G



```

<400> 713
nntggctgcc tgngcgtnta ctctaaagga tntactatnc atatggantc naanacgact 60
cactacacgg cncctcncgg agccnnggtc agtgccctnct nggagacctt ctctggggca 120
ggangagcac tnggtatgtt cacgtatcnc ttcntaaana tacnnccctc cg 172

```

```

<210> 714
<211> 112
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(714)
<223> n=A,T,C or G

```

```

<400> 714
nttgcgtgcc tggacgtnta ctctgcanga tctactactc atgngaattc taantacgga 60
ctcactatnc ggcancgcag gcgcagcagg gaanggggtca cctcccagtc tc 112

```

```

<210> 715
<211> 326
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(326)
<223> n=A,T,C or G

```

```

<400> 715
tactctanag gatctncgng tcatntggat tctatntcga ctcaactctag ggctcnagcn 60
gtcngccggg caagttattc ggatcgtcgg gntccgagct tcgcaattaa ntgtgccatc 120
gttctncaac gttcctgact nggaancccc ngcngttcng atccncnggt acctagctcc 180
anntcccccg tntccttctt ggngtntcat naangaggac cncctctgat cnccttctct 240
taatctgcnc acnctgaacg nccaatggac atngtgcggt taatntanna ggcccgnntc 300
gngtgccctt cccgtnannt cagctc 326

```

```

<210> 716
<211> 122
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(122)
<223> n=A,T,C or G

```

```

<400> 716
nntgcgtcgc ctgngcgtnt actctagatg atctgantag tcatatggat tctaatacga 60
ctcannatag ggctctagcg nggatncnga ttctgctntcc ngattcantg acnccggtan 120

```



122

```
<220>  
<221> misc_feature  
<222> (1)...(203)  
<223> n=A,T,C or G
```

```
<210> 718
<211> 168
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(168)  
<223> n=A,T,C or G
```

<400> 718

ggcagganga	tcncttgagc	ccnngaggtc	gaggctacag	tgagccanga	gtgcactact	60
gtnnccgcct	ccgcatnca	gngtggtccg	atccccgggt	accganctng	anttcaactgg	120
anttcttttt	aancgtnttg	antggtacna	ccctcgantc	cctggctg		168

```
<210> 719
<211> 210
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(210)  
<223> n=A,T,C or G
```

```
<400> 719
cancgtcgnc ataacacgta tttntgatn aagattctna ctgacccatn aantctacnt 60
ctcaagctct tncanngtcc agtnaangga atgtgtatnn gtngggatnc cacanaaaaa 120
aganatntcg gncgcttcat tantcctcct tcttaccan ntctctngat nncagntntg 180
ancntgaacg cacactacng gatntctcca
210
```

$$\begin{array}{ll} \langle 210 \rangle & 720 \\ \langle 211 \rangle & 131 \end{array}$$



<220>



<223> n=A,T,C or G

```
cctccggaat atccaantag agtaantnct ctctaattcg gggnaattgg nggggttnnat 60
acgtcctcct cccccagnt aggattnana aaaggntctc cagancaaaa nctccaaagt 120
gnatcnanta qccqtncccg anattcaacg cccctacgtc          160
```

<213> Homo sapiens

<223> n=A, T, C or G

```
tnanccnata  tacaccaaata  tctgattcta  aantcccacc  caagggaaaa  aagttgagaa  60
gagccttttc  actttttctac  taataaaaaa  atgcaccagc  ccctaccann  agtgnggaaa  120
acctccttag  gcccttgntt  ggaacaancc  aaaatc      156
```

<213> Homo sapiens

<223> n=A,T,C or G

aganggttnt	atncatgctg	tactcgcgcg	cctgcagtcg	acactagtgg	atccaaagaa	60
ttcggcacga	gagacggtgc	gcatggaacc	gagggcccca	gccgngagg	cgcgcgcgcc	120
gagcccgcgg	ncagacgccc	catcagtagc	gtccgcaccg	ggnagccgcg	gntctcgccc	180
gagccgtggg	cgcgcgcgag	gggggggctc	gcctcccgcc	gtccctcgca	gctctgcggg	240
gcccagagccc	gcgcgcgtcg	cgcgcgcgnc	ttgcgcgtcg	gnccgcgcgg	nccggnaaac	300
gcggtcgagg	tctggatgng	gcanngcccg	cnctntcgcg	tgagcct		347

<213> Homo sapiens

<223> n=A,T,C or G



```
<400> 729
cngactgctn gcgtttaaac ttaagcnagg taccgaacgg ggatnnacga ctantgatcg 60
gctggctgct tccagtcgat tanatttgtg aaaaagctga accncngccn gttaaggggg 120
annatgcaaa anatncatcc nntgccccn taaactgntc tntccnaggg aaaaaangga 180
ag 182
```



<400> 732  
gcttggtacc gagctnggat ccctagtaac ggcgcgcagt gtgctggaat tcggctttct 60







```
<210> 735
<211> 126
<212> DNA
<213> Homo sapiens
```

```
<400> 735
ncnttgaaac nggttgacca gacttcaggc ctgtgcgctc aatcgtggag aatctcgtgc 60
cgaattcggc acgagtctct ctctctctct ctctctctct ctctctctct ntctctctct 120
ctctctct                                     126
```

```
<220>  
<221> misc_feature  
<222> (1) ... (165)  
<223> n=A,T,C or G
```

```
<210> 737
<211> 125
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(125)  
<223> n=A,T,C or G
```



```
<210> 738
<211> 137
<212> DNA
<213> Homo sapiens
```

```
<400> 738
ggagncnctt gancaggatg accgacttca ggctgtgtcg ctcaatcgtg gagaatctcg 60
tgccgaattc ggcacgagtc tctctctctc tctctctctc tctctctctc tctctctctc 120
tctctctctc tctctctc                                     137
```

```
<210> 739
<211> 970
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(970)  
<223> n=A,T,C or G
```

```
<210> 740
<211> 739
```



<213> Homo sapiens

<221> misc feature

<223> n=A,T,C or G

gntgtcnaaa	aagcaggctg	gtaccggctc	ggaattcgcg	gccgcgtcga	cggcccttgg	60
tgccactagt	tctttcattc	ttccccccca	tcaatcagtg	aacttttttag	cctaactcaa	120
gctttgctcc	aatgcatagg	atztatgatt	gtggggattt	ccagataata	taaataattca	180
acatgaatat	tttaaattaa	ggcatgagac	atttttccta	actgagcata	gccatgaacc	240
tctcacgtct	gttcctctgt	gncagtttgt	agcactgaat	acagcagccc	tcctaaaagt	300
ccaggcagtg	cacaggctct	gacatgatga	agtgacgtgt	tgctatggtg	attttgcagc	360
tggccaaata	gtcactgggt	gattttaccc	agcaggagat	ttttgcaaaa	atttcctggg	420
tgagagtga	atcaaactcc	tattttgttt	ctcctctgca	agctgnagtt	aanatggatt	480
aatgagtact	tttagattaa	ttaactctga	agagaaaatg	ggagaaaagn	gaggaaggtt	540
gttggcagaa	gtcattgctg	gaatccttct	gaaggagata	ctgacttcac	ttgcaaagac	600
aagagactan	aagacaatga	agttaaactt	ggcctgtctn	tcatatgata	gatgcttgag	660
agtacaggnt	cagggaaatt	ttaattctgn	catacgcata	ttggattatg	tgggtcatgg	720
ctttgttttg	cncctaacc					739

<211> 1171

<213> Homo sapiens

<221> misc feature

<223> n=A, T, C or G

gccttgnnggt	gacactatag	aacatgtttg	tacaaaaaag	caggctggta	cgggtccgga	60
attcgcgggc	gcgtcgacgg	cccttnntgc	cactagtctc	ttcattcttc	ccccccatca	120
atcagtgaac	tttttagcct	actcaaagct	ttgctccaat	gcataggatt	tatgattgtg	180
gggatttcca	gataatataa	atattcaaca	tgaatatttt	aaattaaggc	atgagacatt	240
tttcctaact	gagcatagcc	atgaacctct	cacgtctgtt	cctctgtgtc	agtttgtagc	300
actgaataca	gcagccctcc	taaaagtcca	ggcagtgcac	aggctctgac	atgatgaagt	360
gacgtgttgc	tatggtgatt	ttgcagctgg	ccaaatagtc	actggttgat	tttaccagc	420
aggagatttt	tgcaaaaatt	tcctgggtga	gagtgaaatc	aaactcctat	tttgtttctc	480
ctctgcaagc	tgtagttaag	aagggtattaa	tggagtactt	tttaagaatt	aaattaacct	540
cttgaaagaa	gaaaaaatgg	gggaagaaaa	aaagtggaag	ggaaaagggg	ttggttttgg	600
gccnaaaaaa	aagttccaan	tttnggcntt	ggggaaaaat	tccccntttt	ccttggnaaa	660
aggggggnaa	ggttaancct	tgggaacctt	tttccnncct	tttnggcccc	aaaggggaac	720
ccanggggaa	agaaccttta	ggnaaaggaa	accattttgg	gaanggggtt	naaaacctnt	780
ngggcccccg	ggccctcctc	caanaaggga	aaaaaaaagg	cctggaaaaa	gtaccagggt	840
ttcangggna	aaanttaaaa	ttcttggcca	atanncctat	aattgggaat	tatggggggg	900
ccatgggctt	ttggtttggg	cnccttaacc	cgcnttttaa	attcaanna	aaaaaaagng	960



```
<210> 742
<211> 739
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(739)  
<223> n=A,T,C or G
```

<400>	742					
gntgtcnaaa	aagcaggctg	gtaccgggtcc	ggaatttcgcg	gccgcgctcga	cggcccttgg	60
tgccactagt	tctttcattc	ttcccncca	tcaatcagtg	aacttttttag	cctaactcaa	120
gctttgctcc	aatgcatagg	atztatgatt	gtggggattt	ccagataata	taaataattca	180
acatgaatat	tttaaattaa	ggcatgagac	atttttccta	actgagcata	gccatgaacc	240
tctcacgtct	gttcctctgt	gncagtttgt	agcactgaat	acagcagccc	tcctaaaagt	300
ccaggcagtg	cacaggctct	gacatgatga	agtgacgtgt	tgctatggtg	attttgcagc	360
tgccaaaata	gtcactgggt	gattttaccc	agcaggagat	ttttgcaaaa	atttcctggg	420
tgagagtgaa	atcaaactcc	tattttgttt	ctcctctgca	agctgnagtt	aanatggatt	480
aatgagtact	tttagattaa	ttaactctga	agagaaaatg	ggagaaaagn	gaggaagggt	540
gttggcagaa	gtcattgctg	gaatccttct	gaaggagata	ctgacttcac	ttgcaaagac	600
aagagactan	aagacaatga	agttaaactt	ggcctgtctn	tcatatgata	gatgcttgag	660
agtacaggnt	cagggaaatt	ttaattctgn	catacgcata	ttggattatg	tgggtcatgg	720
ctttgtttgg	cncctaacc					739

```
<210> 743
<211> 610
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(610)  
<223> n=A,T,C or G
```

<400> 743						
ctgtccttat	ttcttttagca	aaaattttccc	aagagaagaa	ttgctgggat	aatgcacatt	60
taaattttttg	atagacattc	ccaaatatta	tacctgtttt	tgagaccttt	aattcctgtt	120
gtcaaattgc	cctatatatg	gagtaataaa	cacgatttaa	agaaatgagg	actaaaaaaaa	180
gattatatat	aacccaacat	aaaggcaacc	tcttaggcgt	tgacagaaac	tgacaacttt	240
ttatctgtgg	gtgcgatcca	ttataagtaa	cctgagcacc	ttattttttc	tttttaaact	300
ctaggtagga	tacccgaggt	ccacaaaattt	ttcataagaa	atattttttc	tctgccctat	360
gagatttttaa	aaaatattat	actgcttcaa	ttgcatcaaa	agaaatggac	cctaatactc	420
atgatgaag	atttggaggt	agaagaccctg	agttttcaatt	ttggcatggc	tgtttgtcta	480
gctctgnagat	cttggacagg	tcaattgact	tggcttaatc	ttctcatcca	tttagngggag	540
acagcacac	tattcacagg	actattgncn	gaattaccag	acaatagcat	aggngaaaaat	600







```
<210> 747
<211> 738
<212> DNA
<213> Homo sapiens
```

<400>	747						
gatatcccg	gaattcg	ccgcgtcnac	gaagcacaga	cctgngccct	gctctcatgg	60	
ggcagactgc	catttgtcat	tnattactga	aggaaaggga	tcctcagttt	gcttgtggac	120	
atttcaaatt	tgaggtgaga	gttgataag	taagaataaa	gctgctcttc	aaagagatga	180	
atatagaaaa	agaaacaaga	tacagncttg	gcagtaaggc	tgggaggaag	gggaaaagg	240	
aataaagaat	gaaagagtga	gaaatgtgag	caggagctga	acacagaaaa	gttcagngac	300	
agaagcanaa	ggagggaaga	agggaggagg	gtccctttca	cagaggetca	cgaggatgct	360	
ttatgngtgc	catgcagtcc	atgttcagga	tgtctgcttc	ttanctctct	acttttctaa	420	
tanaaatttg	gatacttact	gaccta	atgtaacagg	gagagaagg	gaatttcaa	480	
gcantaaatt	gaaaaattgt	tcacaatttc	atTTTTTaa	aaaaggagg	taacagaaga	540	
agaggttaat	gtggtaat	taggatgnct	cttgcgacac	atgaatgnat	ctggtatcat	600	
ctgagtggga	ggggagctgt	cttctgacc	caaaggatc	ctttcgttan	ccngnaetta	660	
ngtcccaaaa	cctcaccacc	ttggagaaat	natttccttt	tgggggtntc	attaaanct	720	
tttggncccc	gcaaaagc					738	

```
<210> 748
<211> 647
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(647)  
<223> n=A,T,C or G
```

<400> 748  
ctntgtggcg gtggctgtct catttggttg gacttttttg gtcgtaggaa cctggatatng 60  
aggtcgagag taagacgggc tattagtagt cgcacggag ttattttgtga aaacctggtt 120



```
<210> 749
<211> 642
<212> DNA
<213> Homo sapiens
```

<400> 749							
ctntgtggcg	gtgngtgtct	catttggggtg	gacttttttg	gtcgtaggaa	cctgggtatgc	60	
aggtccgcgg	agcgtgggct	ctcgtcgtgg	atgttggggg	ttgggtgtggt	gocgggttggt	120	
tttggttctg	ttgagcgtag	tgtgtttgaa	ggttagcgtt	cgtgtccttc	ttgtgggtttg	180	
gtgtttaggg	cgggtgggga	ggttgtttgt	tagctgttgt	atgtcatatt	gttgggtgttg	240	
ctgccctgtg	ctgtttgtcc	ttgggtattg	tggttgttac	ccgcgcctgtg	tggaaagtgtt	300	
gtggcagggc	gggaatttaa	gtgggagagt	tgtgggacct	gtggttggtg	ttacgtttgct	360	
gcttttgtcg	tgggcggtgg	cggcgcgctct	gataattaga	attggatacg	gagtggtataa	420	
tacttctagt	aaatggggac	ctagtgcctg	acttcccgga	atagggatct	atgcgaagtc	480	
cttaggatag	tctttgataa	gtttaacgcc	cacgacctta	aaattataca	cgatttagacg	540	
cataacgact	cctccaggaa	agataaagaa	tctcacatat	agaacgggac	cccatacacg	600	
tcggaatagg	aacaagaqaa	ctaatttttng	ttaaaaagac	tt		642	

```
<220>
<221> misc_feature
<222> (1)...(639)
<223> n=A,T,C or G
```

<400> 750						
tttgtggcgg	tggtgtctca	tttgggtgga	tttttgggtc	gtaggtaacc	tggtatngag	60
gtatagatgc	cgattggtcc	cgacgagcgt	cacgataaat	tcggtagttt	cgcccttttt	120
agaaggcgct	agtactcgga	acttcacttc	atctcggtag	tttacttttg	cgtatatagc	180
cttctccctc	gaagactagc	cgtcacattc	gttccctag	aatcgtttct	gcccctaaga	240
atccgagagc	gagatccga	aactagagga	accttagaag	agtcgtattt	ccacaaggac	300
cccacagtca	ttccgggaaa	atccctagga	ccatacgtt	aggattcccc	cggaaaccgg	360
aqcaaagctc	atgatttccc	acaccgcgag	agcgcctata	accctatccc	atttcttcgg	420



```
<210> 751
<211> 637
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(637)  
<223> n=A,T,C or G
```

```
<210> 752
<211> 644
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<222> (1)...(644)
<223> n=A,T,C or G
```

<400>	752					
tntgtggcgg	tggtgctcat	ttgggtggat	ttttgggtcg	taggaacctg	gtatgaggtc	60
ttgcgagttg	ttggtgtgtc	ctgtcgttcg	gtggttcctt	tttgagttga	gtttgtcctt	120
tgaggttggt	agctgctggt	cgtttgtgtt	cgtgtagtgc	tttgggttga	gagggttatg	180
gtggtggtta	cgggtgtattg	tcgcccgtgg	tcgcgggggt	ggggtggtcg	tcggttttgt	240
ggttcatagt	agtcttctgc	gttcggtggg	gcgggtttgg	gtgagtagtt	tcgttcttgg	300
atgtcccatt	gacccgccat	aatctaagta	agggttagta	gaaacctctc	cccgatagac	360
acaaccgtcg	tccactaaag	acctcgcttc	tgatttttaa	aaggaccoga	aaaacatccc	420
ttcaacggaa	aaaacggaaa	aaaagtcagc	gaattcaaag	aagccacggg	agagaaaaaa	480
gaactaaagt	tagtccgtca	ttatatgtct	cctcggagga	ggaagcggcg	gtggcgga	540
atgaggcggg	aagaaagacg	acctctatcg	gcggcttang	ccctaaaagg	gcgatacctt	600
acgggatgat	aaggacccta	ggacgcctcc	ttctcggatc	gtcc		644



```
<220>  
<221> misc_feature  
<222> (1)...(635)  
<223> n=A,T,C or G
```

```
<210> 754
<211> 721
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(721)  
<223> n=A,T,C or G
```

<210>	755
<211>	721
<212>	DNA



<211> 782



<213> Homo sapiens

<221> misc feature

<223> n=A,T,C or G

ggccctcgag	gggatactct	agagcggccg	ccgactagt	agctcgtcga	cgatatcccg	60
ggatttgaga	ccaggagaca	gctccagatg	ctgtcagccc	agtgcctggg	gcaggcttcc	120
atctgtgaag	tggagaggcg	ctttgggctt	cttcgttggc	atcaggtgcc	catacctagg	180
gcagctgtgg	aagtgtcagc	gtcctccctg	agaggaactc	ctgctccggg	ggctcctcag	240
tccttcctgc	agtatgctgt	aaagcaccca	catggtaatg	gggngggact	ggtagcatga	300
ctgntccctt	aaaagggtgg	cttccnaag	aaaggagaat	tcttggacna	gggatttcac	360
ttgnttagaa	atgggaaaaa	ttaccatta	gaattttcgn	ttccaaggcn	tnaagncccta	420
aaaggccttt	gattcccgaa	ccttaaccct	gggcagttaa	cctttcaaac	gggataaacc	480
ctgangggga	aaatnaaatc	ctttaaaaaa	gggggggttt	naaggagggc	tctttggctt	540
tcaggcantt	gccaacctgg	gaaattcana	ggggaagtnt	ttttttttgc	ctgcctaggg	600
aacctttact	taaacnaacc	cttgncctcc	catttggggt	tgactttcan	cctaattgct	660
gaaaggaccg	ggcgnntttt	gntttccttt	gncccaaagg	naaanaaacg	gggtgccantt	720
cccangggat	tanttcocga	aaatttggnn	aatttttntt	tgnaactttt	tgggtttttt	780
cc						782

<211> 647

<213> Homo sapiens

<221> misc feature

<223> n=A,T,C or G

ntttgtggcg	gtggtgtctc	atttgggtgg	actttttggg	tctaggaac	ctggtatnga	60
gggaagagcg	ccgtcggtc	gagtacagta	tggagtagta	tagtcttcgc	gccttctcgg	120
gcg'gcggg	tattctctcc	aaaggcagag	gtccctagtc	gacctcgctc	ccctagggtta	180
ggaacagccg	togaatattt	taggttcgtc	gaggctttct	tccgagctct	acgcctaagt	240
agctccgcga	gcaaagtatc	ggtcattttc	ccctatccat	cactccocta	agtacgcctc	300
attattccgg	aaggcaagag	gccagcattc	ctccttagag	tagagggtag	gtacctccgt	360
cgcgtgccgc	gaaagggcag	agcttcgtgt	cttccctccg	cagcagctta	acggtctacg	420
taggcgtttc	cgatcttttc	acgggaatcg	gggtccggga	gggcggcgga	aaacgtcgac	480
gtctcggtea	ccgtcaccgc	cccgaacaac	tagcggcttt	cgcgtttcaa	ctgaggaacc	540
ccgcacccct	cattagcgct	tacgaataatc	gggagtgat	tgcgccaaat	cgttagcctt	600
cqataattat	tctctattag	cqgtccatc	tcgcgctttc	qatttat		647

<211> 657

<213> Homo sapiens



<223> n=A,T,C or G

ctttgtggcg	gtggtgtctc	atttgggttg	acttttggg	tcgtaggaa	ctggtatnga	60
gggctctata	gaaagcctct	tgtctttaga	tacgggcttt	ctggctcctc	gttctggaag	120
tgtagtagta	ggtactgcgg	gaaggcgaag	agtcctttca	aggacgattt	acttaagttg	180
gcttattcta	tagttccttc	gggacataag	gtcgttacga	tctatactgc	gtgggaagct	240
gataggttgg	gacttaaggc	gaataagaag	gaggcggcgg	aggtcgcgat	taccgcagag	300
atattattta	cggcggccgc	gggtaccgcg	ggtcatgcgg	aaattttctg	aggttcttgg	360
attcctaaga	tcgtcccgt	cgagtatact	agcgacgaac	gtaagagtgc	cctcacaaga	420
accggtacaa	actcaagaag	aagttcccat	taagcatcgt	aagaaccggt	aggacagga	480
cggtaagaag	taatcggaga	aaggatccta	ctngttacga	agaagcatcg	ttnagctact	540
ttgcgctacc	gtttatatatt	agacgtgttc	cgtccttctc	cgtgtttana	aaaaaggttt	600
attccgacgg	gagacttagg	cgaatggagg	gttcgcgcgt	tganaatcgg	ancgggg	657

<213> Homo sapiens

<223> n=A,T,C or G

ctttgtggcg	gtggtgtctc	atttgggtgg	actttttggg	tcgtaggaac	ctggtatgna	60
ggaaaagaag	taagcctcga	agcctatctc	cgaccgtatt	tatttcgcag	aagacggaac	120
tacggacgtc	gttaaccccc	agtagccccc	gtaagaaaag	actaaagcga	atggaaaagt	180
cgggaattcc	ggcggagggg	cggcgattac	tgaaaggagt	aagagtaaga	ctattgcgat	240
acttgaggcg	ttccctctta	aaaggcaccc	gaaacactct	attaaaaaac	acccgaagaa	300
gaacaactca	tgcgatcggc	cgtgtgcagc	cgtcaatagt	aaagagagcc	atgaaccatg	360
ccatccttag	accaattagg	atgaagaaga	ggaggaagat	gaggacccaa	ccctaccac	420
tcggaatacc	ccgcacgagc	ctccgaacaa	aatccgggaa	ttaaaacggc	ggccccattc	480
cgcactctcg	tagecgggac	cgaatagaaa	accggaacct	acagctaaag	ggtcctttcc	540
cgctctgttat	ctaccacccc	gcaatccgat	ctccccccc	cctcgtccaa	aaaccctaac	600
ctctgcggca	acattagagc	agaaggagag	ggcgatccct	tgan		644

<213> Homo sapiens

<223> n=A,T,C or G



```
<210> 762
<211> 628
<212> DNA
<213> Homo sapiens
```

[illegible]

```
<210> 763
<211> 147
<212> DNA
<213> Homo sapiens
```

```
<220>  
<221> misc_feature  
<222> (1)...(147)  
<223> n=A,T,C or G
```

```
<400> 763
cattgtgttg gggcagagat aaataattcc tctgaaaagt gttttattgg aatttcaaat 60
gaaaagctaa ctggataact tacagcatgt ttctgccaat aatctcttan aacaggcctc 120
ttttttttat qcacaccacc ttcnggc                                     147
```



<210> 764  
 <211> 146  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(146)  
 <223> n=A,T,C or G

<400> 764  
 cattgtgttg ggtatgtttt ttgaaggcag gtggacagga tttgctgatg ggtaaattggc 60  
 agagtttaggg ggactgttag aacagagaaa ganatcatgg gggtgggttt gactctgatg 120  
 nnnaactggt gccgnntgct cagtat 146

<210> 765  
 <211> 129  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(129)  
 <223> n=A,T,C or G

<400> 765  
 tncncgattc gntnctagcg tntacactna tgtcttggtta ccgagctcgg atccactagt 60  
 ccagtgtggg nggaattcca ttgtgttggg gcaggaggng ctttgngtac ngtgcggctg 120  
 nagaggcgg 129

<210> 766  
 <211> 175  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <222> (1)...(175)  
 <223> n=A,T,C or G

<400> 766  
 cattgtgttg ggcctagtc gaatactttt agtaacttca gacagatctc ctcatctctt 60  
 tctggggctt ggnttttctc ctttgtanaa tgatgccttt ctgtgggttt gtcatttcta 120  
 acattctgtg ngtgatgagg tgtatattcg angantcta tcnccanagt actct 175

<210> 767  
 <211> 602  
 <212> DNA  
 <213> Homo sapiens



<220>

<221> misc\_feature

<222> (1)...(602)

<223> n=A,T,C or G

<400> 767

```
nnnttttaaaa nctgtntctcc ccgcgggtggc ggccgctcta gaactagtgg atccttttcca 60
cctgggtttgt tttcagtgtt taatcctatt agtatcagca ggatataggt caggatatca 120
ggtgcagaac ctgtggaatc agccaatttg gcttgctcat ttactttaat aagggtcccat 180
aatgagttag agtacaaagt tcaagccctg ttgaggggtct gcattaaact ctcagaagta 240
tttagagtgt gccaggagcc gcgaagggtct gggttcgggtg gtggcgggaa ctgtattaga 300
gtgctaggca cggcgcgaca aagtctgtcc aacccaaaac ggtgctgagg cgttgggtgt 360
gagctccagt actcagaaaa gcctctcagc aggtactcaa cagatcctca ggggcttggg 420
ggcccagcac tggcagttag ggcatgaaag acataaaagg gcactacctg tgggtatttt 480
ctgttctcca aggaggaagt agcaaaaatt aggacgtgg aatatactat gttgtagcaa 540
tcccagaaca actgatgctc aacaaatacc acacaaaaca aattttttta aatttaattct 600
ta 602
```

<210> 768

<211> 671

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(671)

<223> n=A,T,C or G

<400> 768

```
tccaccgcgg tggcgggcgc tctagactag tggatccact agtccagtgt ggggtgggaat 60
tcgcggcncg cgtcgacaaa aatactgcta aagtaatatt tttatagatg actatttgcc 120
ttggggccag gaaaagcagc tggagttatt cacttagtac catttttaca tactaacttt 180
gccttttcca tgcttgcttg atgcggcttg cagcactgaa gaacagtttc aattgctagc 240
caaccagaga gcatgatcaa accaaacaag ttccctgttt caggaaaaac aggttttagg 300
taactgaagg gttaccagtt actgattcca caatcttctc tgtaaaanatt ttctgcctat 360
tatgcagact gggcggtctt aaanntggta aaactatnaa ataccatac aatattttta 420
nggggccccn ttatnaagct tttcaggcct tcccctttcc atagcattgg tgggatacaa 480
gaaaccttta aacagcaacn agctatcnag gcccaaaagg aaagtaattn tgatttttta 540
nagattccgn aacgaaaaaa tggctgggtt caaatnacn cttcttttta aaatggnttc 600
cttattaaac nttttttttt ttttaatttta ccccatggtc ntgatnttng ngcttccgcc 660
canaaaatng n 671
```

<210> 769

<211> 877

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(877)

<223> n=A,T,C or G







<221> misc\_feature

<222> (1)...(156)

<223> n=A,T,C or G

<400> 771

```

ttaaaaaanct ggnctccccg cggtggcggc cgctctagaa ctagtggatc cactagtcca 60
gtgtggtgga attcgcgcc gcgcgaccg cgagcggtcg cccctttttt ttttttttn 120
ngtttttttg aanaattcat tgggtattta ttattc 156

```

<210> 772

<211> 586

<212> DNA

<213> Homo sapiens

<220>

<221> misc\_feature

<222> (1)...(586)

<223> n=A,T,C or G

<400> 772

```

ncaanctggn ctccaccgcg gtggcgcgcg ctctagacta gtggatccac tagtccagtg 60
tggtggaatt cgcgcgcgcg tgcgtcacaa agtgcacaca agtcnngnat ttattttatc 120
tccagatatg aaacttaccc ccagctatgg tcttctattt gttattttaat ttctaggcca 180
atTTTTTcca cttgaatgtc agtatTTTaa ttcaaagtca ccttgtccaa ataccaagtc 240
atcaacttac cctcaaatta taccctcatt cagaaaatct acatctatta atggtagcta 300
ttttatccct gccccctgot ttttcttttt atattttaatt aatttgntca tccagcaaat 360
gcttattgag caggtattgt aggcataaaca attctanact ttaaggggac acagnttgca 420
aaacaaaatc ctgccttgna tggatactta tgnnatggng ggatacagac aatcaacata 480
atgangngca tcatatataa tggttagnan aatgataagg gnttttggga aaaaaatgca 540
cccanccaan anggattggg aagtggangg ganggtcang ggangg 586

```

<210> 773

<211> 2983

<212> DNA

<213> Homo sapiens

<400> 773

```

agagatagag tcttccttgg cattgcagga gagaatctga agggatgatg gatgcatcaa 60
aagagctgca agttctccac attgacttct tgaatcagga caacgccgtt tctcaccaca 120
catgggagtt ccaaacgagc agtctgtgtg tccggcgagg acaggtgttt cacctgcggc 180
tggtgctgaa ccagccccta caatcctacc accaactgaa actggaattc agcacagggc 240
cgaatcctag catcgccaaa cacaccctgg tgggtgctcg cccgaggacg ccctcagacc 300
actacaactg gcaggcaacc cttcaaaatg agtctggcaa agaggtcaca gtggctgtca 360
ccagttcccc caatgccatc ctgggcaagt accaactaaa cgtgaaaact ggaaaccaca 420
tccttaagtc tgaagaaaac atcctatacc ttctcttcaa cccatggtgt aaagaggaca 480
tggttttcat gcctgatgag gacgagcgca aagagtacat cctcaatgac acgggctgcc 540
attacgtggg ggctgccaga agtatcaaat gcaaaccctg gaactttggt cagtttgaga 600
aaaatgtcct ggactgctgc atttccttgc tgactgagag ctccctcaag cccacagata 660
ggagggaccc cgtgctggtg tgcagggcca tgtgtgctat gatgagcttt gagaaaggcc 720
agggcgtgct cattgggaat tggactgggg actatgaagg tggcacagcc ccatacaagt 780

```



```

ggacaggcag tgccccgatc ctgcagcagt actacaacac gaagcaggct gtgtgctttg 840
gccagtgtgt ggtgtttgtt gggatcctga ctacagtgtt gagagcgttg ggcattcccag 900
cacgcagtgt gacaggcttc gattcagctc acgacacaga aaggaacctc acggtggaca 960
cctatgtgaa tgagaatggc aagaaaatca ccagtatgac ccacgactct gtctggaatt 1020
tccatgtgtg gacggatgcc tggatgaagc gaccggatct gcccaagggc tacgacggct 1080
ggcaggctgt ggacgcaacg ccgcaggagc gaagccaggg tgtcttctgc tgtgggcat 1140
caccactgac cgccatccgc aaagggtgaca tctttattgt ctatgacacc agattcgtct 1200
tctcagaagt gaatggtgac aggtcatctt ggttgggtgaa gatggtgaat gggcaggagg 1260
agttacacgt aatttcaatg gagaccacaa gcatcgggaa aaacatcagc accaaggcag 1320
tgggccaaga caggcggaga gatatacct atgagtacaa gtatccagaa ggctcctctg 1380
aggagaggca ggtcatggat catgccttcc tcttctcag ttctgagagg gagcacagac 1440
gacctgtaaa agagaacttt cttcacatgt cggtaacaatc agatgatgtg ctgctgggaa 1500
actctgttaa tttcacctgt attcttaaaa ggaagaccgc tgccctacag aatgtcaaca 1560
tcttgggctc ctttgaacta cagttgtaca ctggcaagaa gatggcaaaa ctgtgtgacc 1620
tcaataagac ctgcagatc caaggccaag tatcagaagt gactctgacc ttggactcca 1680
agacctacat caacagcctg gctatattag atgatgagcc agttatcaga ggtttcatca 1740
ttgcggaaat tgtggagtct aaggaaatca tggcctctga agtattcacg tctttccagt 1800
acctgagtt ctctatagag ttgcctaaca caggcagaat tggccagcta cttgtctgca 1860
attgtatctt caagaatacc ctggccatcc ctttgactga cgtcaagttc tctttggaaa 1920
gcctgggcat ctctcacta cagacctctg acctgggac ggtgcagcct ggtgagacca 1980
tccaatccca aataaaatgc accccaataa aaactggacc caagaaattt atcgtcaagt 2040
taagttccaa acaagtgaag gagattaatg ctcagaagat tgttctcatc accaagtagc 2100
cttgtctgat gctgtggagc cttagttgag atttcagcat ttctacctt gtgcttagct 2160
ttcagattat ggatgattaa atttgatgac ttatatgagg gcagattcaa gagccagcag 2220
gtcaaaaagg ccaacacaac cataagcagc cagaccaca aggccaggctc ctgtgctatc 2280
acagggtcac ctcttttaca gttagaaaca ccagccgagg ccacagaatc ccatcccttt 2340
cctgagtcac ggcccaaaa atcagggcca ccattgtctc aattcaaata catagatttc 2400
gaagccacag agtctctccc tggagcagca gactatgggc agcccagtgct tgccacctgc 2460
tgacgaccct tgagaagctg ccatactctc aggccatggg ttaccagcc ctgaaggcac 2520
ctgtcaactg gagtgtctctc tcagcaactg gatgggctg atagaagtgc attctcctcc 2580
tattgcctcc attctcctct ctctatccct gaaatccagg aagtcctctc cctggtgctc 2640
caagcagttt gaagcccaat ctgcaaggac atttctcaag ggccatgtgg ttttgagac 2700
aaccctgtcc tcaggcctga actcaccata gagaccatg tcagcaaacg gtgaccagca 2760
aatcctcttc cttattctta aagctgcccc ttgggagact ccaggagaa ggcattgctt 2820
cctccctggt gtgaactctt tctttggtat tccatccact atcctggcaa ctcaaggctg 2880
cttctgttaa ctgaagcctg ctcttcttgg ttctgcctc cagagatttg ctcaaagat 2940
caataagctt taaattaaac tctacttcaa gaaaaaaaaa ccg 2983

```

<210> 774

<211> 3064

<212> DNA

<213> Homo sapiens

<400> 774

```

aattctaaaa atgcttttgc aagcttgcac gctgcaggc gcagcggccg ccagtgtgat 60
ggatatctgc agaattcggc ttgcgctcag ctggaattcc gcagagatag agtcttccct 120
ggcattgcag gagagaatct gaagggatga tggatgcac aaaagagctg caagttctcc 180
acattgactt cttgaatcag gacaacgccg tttctcacca cacatgggag ttccaaacga 240
gcagtctgtg gttccggcga ggacagggtg ttacctgcg gctggtgctg aaccagcccc 300
tacaatccta ccaccaactg aaactggaat tcagcacagg gccgaatcct agcatcgcca 360
aacacaccct ggtggtgctc gacccgagga cgccctcaga ccactacaac tggcaggcaa 420

```



```
<210> 775
<211> 684
<212> PRT
<213> Homo sapiens
```







Gly	Ile	Leu	Thr	Thr	Val	Leu	Arg	Ala	Leu	Gly	Ile	Pro	Ala	Arg	Ser
		275				280						285			
Val	Thr	Gly	Phe	Asp	Ser	Ala	His	Asp	Thr	Glu	Arg	Asn	Leu	Thr	Val
		290				295						300			
Asp	Thr	Tyr	Val	Asn	Glu	Asn	Gly	Lys	Lys	Ile	Thr	Ser	Met	Thr	His
305				310						315				320	
Asp	Ser	Val	Trp	Asn	Phe	His	Val	Trp	Thr	Asp	Ala	Trp	Met	Lys	Arg
				325				330						335	
Pro	Asp	Leu	Pro	Lys	Gly	Tyr	Asp	Gly	Trp	Gln	Ala	Val	Asp	Ala	Thr
		340						345				350			
Pro	Gln	Glu	Arg	Ser	Gln	Gly	Val	Phe	Cys	Cys	Gly	Pro	Ser	Pro	Leu
		355				360						365			
Thr	Ala	Ile	Arg	Lys	Gly	Asp	Ile	Phe	Ile	Val	Tyr	Asp	Thr	Arg	Phe
370						375						380			
Val	Phe	Ser	Glu	Val	Asn	Gly	Asp	Arg	Leu	Ile	Trp	Leu	Val	Lys	Met
385				390						395				400	
Val	Asn	Gly	Gln	Glu	Glu	Leu	His	Val	Ile	Ser	Met	Glu	Thr	Thr	Ser
				405				410						415	
Ile	Gly	Lys	Asn	Ile	Ser	Thr	Lys	Ala	Val	Gly	Gln	Asp	Arg	Arg	Arg
		420						425				430			
Asp	Ile	Thr	Tyr	Glu	Tyr	Lys	Tyr	Pro	Glu	Gly	Ser	Ser	Glu	Glu	Arg
435						440						445			
Gln	Val	Met	Asp	His	Ala	Phe	Leu	Leu	Leu	Ser	Ser	Glu	Arg	Glu	His
450						455				460					
Arg	Arg	Pro	Val	Lys	Glu	Asn	Phe	Leu	His	Met	Ser	Val	Gln	Ser	Asp
465				470						475				480	
Asp	Val	Leu	Leu	Gly	Asn	Ser	Val	Asn	Phe	Thr	Val	Ile	Leu	Lys	Arg
				485				490						495	
Lys	Thr	Ala	Ala	Leu	Gln	Asn	Val	Asn	Ile	Leu	Gly	Ser	Phe	Glu	Leu
		500						505				510			
Gln	Leu	Tyr	Thr	Gly	Lys	Lys	Met	Ala	Lys	Leu	Cys	Asp	Leu	Asn	Lys
515						520						525			
Thr	Ser	Gln	Ile	Gln	Gly	Gln	Val	Ser	Glu	Val	Thr	Leu	Thr	Leu	Asp
530						535				540					



Ser	Lys	Thr	Tyr	Ile	Asn	Ser	Leu	Ala	Ile	Leu	Asp	Asp	Glu	Pro	Val
545					550					555					560
Ile	Arg	Gly	Phe	Ile	Ile	Ala	Glu	Ile	Val	Glu	Ser	Lys	Glu	Ile	Met
				565					570					575	

Ala Ser Glu Val Phe Thr Ser Phe Gln Tyr Pro Glu Phe Ser Ile Glu  
580 585 590

Leu Pro Asn Thr Gly Arg Ile Gly Gln Leu Leu Val Cys Asn Cys Ile  
595 600 605

Phe Lys Asn Thr Leu Ala Ile Pro Leu Thr Asp Val Lys Phe Ser Leu  
610 615 620

Glu Ser Leu Gly Ile Ser Ser Leu Gln Thr Ser Asp His Gly Thr Val  
625 630 635 640

Gln Pro Gly Glu Thr Ile Gln Ser Gln Ile Lys Cys Thr Pro Ile Lys  
645 650 655

Thr Gly Pro Lys Lys Phe Ile Val Lys Leu Ser Ser Lys Gln Val Lys  
660 665 670

Glu Ile Asn Ala Gln Lys Ile Val Leu Ile Thr Lys  
675 680

<210> 776

<211> 679

<212> PRT

<213> Homo sapiens

<400> 776

Met Met Asp Ala Ser Lys Glu Leu Gln Val Leu His Ile Asp Phe Leu  
5 10 15

Asn Gln Asp Asn Ala Val Ser His His Thr Trp Glu Phe Gln Thr Ser  
20 25 30

Ser Pro Val Phe Arg Arg Gly Gln Val Phe His Leu Arg Leu Val Leu

Asn Gln Pro Leu Gln Ser Tyr His Gln Leu Lys Leu Glu Phe Ser Thr  
50 55 60

Gly Pro Asn Pro Ser Ile Ala Lys His Thr Leu Val Val Leu Asp Pro  
65 70 75 80

Arg Thr Pro Ser Asp His Tyr Asn Trp Gln Ala Thr Leu Gln Asn Glu  
85 90 95

[illegible]







Gly 370	Asp	Ile	Phe	Ile	Val	Tyr 375	Asp	Thr	Arg	Phe	Val 380	Phe	Ser	Glu	Val
Asn 385	Gly	Asp	Arg	Leu	Ile 390	Trp	Leu	Val	Lys	Met 395	Val	Asn	Gly	Gln	Glu 400
Glu	Leu	His	Val	Ile 405	Ser	Met	Glu	Thr	Thr 410	Ser	Ile	Gly	Lys	Asn 415	Ile
Ser	Thr	Lys	Ala 420	Val	Gly	Gln	Asp 425	Arg	Arg	Arg	Asp	Ile 430	Thr	Tyr	Glu
Tyr	Lys 435	Tyr	Pro	Glu	Gly	Ser 440	Ser	Glu	Glu	Arg	Gln 445	Val	Met	Asp	His
Ala 450	Phe	Leu	Leu	Leu	Ser 455	Ser	Glu	Arg	Glu	His 460	Arg	Gln	Pro	Val	Lys
Glu 465	Asn	Phe	Leu	His 470	Met	Ser	Val	Gln	Ser 475	Asp	Asp	Val	Leu	Leu	Gly 480
Asn	Ser	Val	Asn 485	Phe	Thr	Val	Ile	Leu 490	Lys	Arg	Lys	Thr	Ala 495	Ala	Leu
Gln	Asn	Val	Asn 500	Ile	Leu	Gly	Ser 505	Phe	Glu	Leu	Gln	Leu 510	Tyr	Thr	Gly
Lys	Lys 515	Met	Ala	Lys	Leu	Cys 520	Asp	Leu	Asn	Lys	Thr 525	Ser	Gln	Ile	Gln
Gly 530	Gln	Val	Ser	Glu	Val 535	Thr	Leu	Thr	Leu	Asp 540	Ser	Lys	Thr	Tyr	Ile
Asn 545	Ser	Leu	Ala	Ile 550	Leu	Asp	Asp	Glu	Pro	Val 555	Ile	Arg	Gly	Phe	Ile 560
Ile	Ala	Glu	Ile 565	Val	Glu	Ser	Lys	Glu	Ile 570	Met	Ala	Ser	Glu	Val 575	Phe
Thr	Ser	Asn 580	Gln	Tyr	Pro	Glu	Phe 585	Ser	Ile	Glu	Leu	Pro 590	Asn	Thr	Gly
Arg	Ile 595	Gly	Gln	Leu	Leu	Val 600	Cys	Asn	Cys	Ile 605	Phe	Lys	Asn 610	Thr	Leu
Ala 610	Ile	Pro	Leu	Thr 615	Asp	Val	Lys	Phe	Ser	Leu 620	Glu	Ser	Leu	Gly	Ile
Ser 625	Ser	Leu	Gln	Thr 630	Ser	Asp	His	Gly	Thr 635	Val	Gln	Pro	Gly	Glu	Thr 640



Ile Gln Ser Gln Ile Lys Cys Thr Pro Ile Lys Thr Gly Pro Lys Lys  
 645 650 655

Phe Ile Val Lys Leu Ser Ser Lys Gln Val Lys Glu Ile Asn Ala Gln  
 660 665 670

Lys Ile Val Leu Ile Thr Lys  
 675

<210> 777

<211> 5668

<212> DNA

<213> Homo sapiens

<400> 777

```

gtcacttagg aaaaggtgtc ctttcgggca gccgggctca gcatgaggaa cagaaggaat 60
gacactctgg acagcacccg gaccctgtac tccagcgcggt ctcgagcac agacttgtct 120
tacagtgaag gcgacttggg gaattttatt caagcaaatt ttaagaaacg agaatgtgtc 180
ttctttacca aagattccaa ggccacggag aatgtgtgca agtgtggcta tgcccagagc 240
cagcacatgg aaggcaccca gatcaacca agtgagaaat ggaactacaa gaaacacacc 300
aaggaatttc ctaccgacgc ctttggggat attcagtttg agacactggg gaagaaaggg 360
aagtatatac gtctgtcctg cgacacggac gcggaaatcc ttacagagct gctgaccag 420
cactggcacc tgaaaacacc caacctgggc atttctgtga cggggggcgc caagaacttc 480
gccctgaagc cgcgcatgcg caagatcttc agcgggctca tctacatcgc gcagtccaaa 540
gggtgcttga ttctcacggg aggcacccat tatggcctga cgaagtacat cggggagggtg 600
gtgagagata acaccatcag caggagttca gaggagaata ttgtggccat tggcatagca 660
gcttggggca tgggtctcaa ccgggacacc ctcatcagga attgcgatgc tgagggctat 720
tttttagccc agtaccttat ggatgacttc acaagggatc cactgtatat cctggacaac 780
aaccacacac atttgctgct cgtggacaat ggctgtcatg gacatcccac tgtcgaagca 840
aagctccgga atcagctaga gaagcatatc tctgagcgca ctattcaaga ttccaactat 900
gggtggcaaga tccccattgt gtgttttgcc caaggagggtg gaaaagagac tttgaaagcc 960
atcaatacct ccatcaaaaa taaaattcct tgtgtggtgg tgggaaggctc gggccggatc 1020
gctgatgtga tcgctagcct ggtggagggtg gaggatgccc cgacatcttc tgccgtcaag 1080
gagaagctgg tgcgcttttt accccgcacg gtgtcccggc tgtctgagga ggagactgag 1140
agttggatca aatggctcaa agaaattctc gaatgttctc acctattaac agttattaaa 1200
atggaagaag ctggggatga aattgtgagc aatgccatct cctacgctct atacaaagcc 1260
ttcagcacca gtgagcaaga caaggataac tggaaatgggc agctgaagct tctgctggag 1320
tggaaccagc tggacttagc caatgatgag attttcacca atgaccgccc atgggagctc 1380
gctgaccttc aagaagtcac gtttacggct ctcataaagg acagacccaa gtttgtccgc 1440
ctctttctgg agaattggtt gaacctacgg aagttttctc cccatgatgt cctcactgaa 1500
ctctttctca accacttcag cacgcttggt tacgggaatc tgcagatcgc caagaattcc 1560
tataatgatg cctcctcac gtttgtctgg aaactggttg cgaacttccg aagaggcttc 1620
cggaaggaag acagaaatgg ccgggacgag atggacatag aactccacga cgtgtctcct 1680
attactcggc accccctgca agctctcttc atctgggcca ttcttcagaa taagaaggaa 1740
ctctccaaag tcatttggga gcagaccagg ggctgcactc tggcagccct gggagccagc 1800
aagcttctga agactctggc caaagtgaag aacgacatca atgctgctgg ggagtcaggag 1860
gagctggcta atgagtacga gaccggggct gttgagctgt tcaactgagt ttacagcagc 1920
gatgaagact tggcagaaca gctgctgggt tattcctgtg aagcttgggg tggaagcaac 1980

```



tgtctggagc	tggcggtgga	ggccacagac	cagcatttca	cgcgccagcc	tggggtccag	2040
aattttcttt	ctaagcaatg	gtatggagag	atttcccag	acaccaagaa	ctggaagatt	2100
atcctgtgtc	tgtttattat	acccttggtg	ggctgtggct	ttgtatcatt	taggaagaaa	2160
cctgtcgaca	agcacaagaa	gctgctttgg	tactatgtgg	cgttcttcac	ctcccccttc	2220
gtggtcttct	cctggaatgt	ggtctttctac	atcgccctcc	tcctgctgtt	tgccctacgtg	2280
ctgctcatgg	atttccattc	ggtgccacac	ccccccgagc	tggtcctgta	ctcgctgggtc	2340
tttgtcctct	tctgtgatga	agtgagacag	tggtacgtaa	atgggggtgaa	ttatttttact	2400
gacctgtgga	atgtgatgga	cacgctgggg	cttttttact	tcatagcagg	aattgtattt	2460
cggctccact	cttcctaata	aagctctttg	tattctggac	gagtcatttt	ctgtctggac	2520
tacattattt	tcactctaag	attgatccac	atttttactg	taagcagaaa	cttaggaccc	2580
aagattataa	tgctgcagag	gatgctgata	gatgtgttct	tcttcctgtt	cctctttgcg	2640
gtgtggatgg	tggcctttgg	cgtggccagg	caagggatcc	ttaggcagaa	tgagcagcgc	2700
tggaggtgga	tattccgttc	ggtcatctac	gagccctacc	tggccatgtt	cggccagggtg	2760
cccagtgacy	tggatggtac	cacgtatgac	tttgccctact	gcaccttcac	tgggaatgag	2820
tccaagccac	tgtgtgtgga	gctggatgag	cacaacctgc	cccggttccc	cgagtggatc	2880
accatcccc	tggtgtgcat	ctacatgtta	tccaccaaca	tcctgctggg	caacctgctg	2940
gtcgccatgt	ttggctacac	ggtgggcacc	gtccaggaga	acaatgacca	ggtctggaag	3000
ttccagaggt	acttctggt	gcaggagtag	tgacgcgcc	tcaatatccc	cttcccccttc	3060
atcgtcttcg	cttacttcta	catggtgggtg	aagaagtgtc	tcaagtgttg	ctgcaaggag	3120
aaaaacatgg	agtcttctgt	ctgctgtttc	aaaaatgaag	acaatgagac	tctggcatgg	3180
gaggggtgtc	tgaaggaaaa	ctaccttgtc	aagatcaaca	caaaagccaa	cgacacctca	3240
gaggaaatga	ggcatcgatt	tagacaactg	gatacaaagc	ttaatgatct	caaggggtctt	3300
ctgaaagaga	ttgctaataa	aatcaaataa	aactgtatga	aactctaata	gagaaaaatc	3360
taattatagc	aagatcatat	taaggaatgc	tgatgaacaa	ttttgtctatc	gactactaaa	3420
tgagagattt	tcagacccct	gggtacatgg	tggtatgattt	taaatcaccc	tagtgtgctg	3480
agaccttgag	aataaaagtg	gtgattgggt	tcatacttga	agacggatat	aaaggaagaa	3540
tatttccttt	atgtgtttct	ccagaatggg	gcctgtttct	ctctgtgtct	caatgcctgg	3600
gactggaggt	tgatagttta	agtgtgttct	taccgcctcc	tttttccttt	aatcttattt	3660
ttgatgaaca	catatatagg	agaacatcta	tcctatgaat	aagaacctgg	tcattgttta	3720
ctcctgtatt	gttattttgt	tcattttcaa	ttgattctct	acttttccct	tttttgtatt	3780
atgtgactaa	ttagttggca	tattgttaaa	agtctctcaa	attaggccag	attctaaaac	3840
atgctgcagc	aagaggaccc	cgctctcttc	aggaaaagtg	ttttcatttc	tcaggatgct	3900
tcttacctgt	cagaggaggt	gacaaggcag	tctcttgctc	tcttggtactc	accaggctcc	3960
tattgaagga	accaccccca	ttcctaataa	tgtgaaaagt	cgcacaaaat	gcaaccttga	4020
aaggcactac	tgactttgtt	cttattggat	actcctctta	tttattattt	ttccattaaa	4080
aataatagct	ggctattata	gaaaatttag	accatacaga	gatgtagaaa	gaacataaat	4140
tgtccccatt	accttaaggt	aatcactgct	aacaatttct	ggatgggtttt	tcaagtctat	4200
tttttttcta	tgtatgtctc	aattctcttt	caaaatttta	cagaatgtta	tcatactaca	4260
tataactttt	ttatgtaaag	tttttcactt	agtattttat	caaatatgtt	tttattatat	4320
tcatagcctt	cttaaacatt	atatcaataa	ttgcataata	ggcaacctct	agcgattacc	4380
ataattttgc	tcattgaagg	ctatctccag	ttgatcattg	ggatgagcat	ctttgtgcat	4440
gaatcctatt	gctgtatttg	ggaaaatttt	ccaaggttag	attccaataa	atatctattt	4500
attattaatat	attaaaaatat	cgattttatta	ttaaaacctat	ttataaggct	ttttcataaa	4560
tgtatagcaa	ataggaatta	ttaacttgag	cataagatat	gagatacatg	aacctgaact	4620
attaaaataa	aatattatat	ttaaccctag	tttaagaaga	agtcaatatg	cttattttaa	4680
tattatggat	ggtgggcaga	tcacttgagg	tcaggagtct	gagaccagcc	tggccaacat	4740
ggcaaaacca	catctctact	aaaaataaaa	aaattagctg	ggtgtgggtgg	tgactcctg	4800
taatcccagc	tactcagaag	gctgaggtag	aagaattgct	ggaacctggg	aggcggaggt	4860
tgacgtgaac	caagattgca	ccactgcact	ccagccgggg	tgacagagtg	agactccgac	4920
tgaaaataaa	taaataaata	aataaataaa	taaataaata	aatattatgg	atgggtgaagg	4980
gaatgggtata	gaattggaga	gattatctta	ctgaacacct	qtaqtcccaq	ctttctctgg	5040



```
<210> 778
<211> 1095
<212> PRT
<213> Homo sapiens
```

Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr

5 10 15

Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu  
20 25 30

Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe  
35 40 45

Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala  
50 55 60

Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp  
65 70 75 80

Asn Tyr Lys Lys His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp  
85 90 95

Ile Gln Phe Glu Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser  
100 105 110

Cys Asp Thr Asp Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp  
115 120 125

His Leu Lys Thr Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys  
130 135 140

Asn Phe Ala Leu Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile  
145 150 155 160

Tyr Ile Ala Gln Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His  
165 170 175











Phe	Phe	Thr	Ser	Pro	Phe	Val	Val	Phe	Ser	Trp	Asn	Val	Val	Phe	Tyr
				725								735			
Ile	Ala	Phe	Leu	Leu	Leu	Phe	Ala	Tyr	Val	Leu	Leu	Met	Asp	Phe	His
				740				745				750			
Ser	Val	Pro	His	Pro	Pro	Glu	Leu	Val	Leu	Tyr	Ser	Leu	Val	Phe	Val
				755				760				765			
Leu	Phe	Cys	Asp	Glu	Val	Arg	Gln	Trp	Tyr	Val	Asn	Gly	Val	Asn	Tyr
				770				775				780			
Phe	Thr	Asp	Leu	Trp	Asn	Val	Met	Asp	Thr	Leu	Gly	Leu	Phe	Tyr	Phe
785				790				795				800			
Ile	Ala	Gly	Ile	Val	Phe	Arg	Leu	His	Ser	Ser	Asn	Lys	Ser	Ser	Leu
				805				810				815			
Tyr	Ser	Gly	Arg	Val	Ile	Phe	Cys	Leu	Asp	Tyr	Ile	Ile	Phe	Thr	Leu
				820				825				830			
Arg	Leu	Ile	His	Ile	Phe	Thr	Val	Ser	Arg	Asn	Leu	Gly	Pro	Lys	Ile
				835				840				845			
Ile	Met	Leu	Gln	Arg	Met	Leu	Ile	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu
				850				855				860			
Phe	Ala	Val	Trp	Met	Val	Ala	Phe	Gly	Val	Ala	Arg	Gln	Gly	Ile	Leu
865				870				875				880			
Arg	Gln	Asn	Glu	Gln	Arg	Trp	Arg	Trp	Ile	Phe	Arg	Ser	Val	Ile	Tyr
				885				890				895			
Glu	Pro	Tyr	Leu	Ala	Met	Phe	Gly	Gln	Val	Pro	Ser	Asp	Val	Asp	Gly
				900				905				910			
Thr	Thr	Tyr	Asp	Phe	Ala	His	Cys	Thr	Phe	Thr	Gly	Asn	Glu	Ser	Lys
				915				920				925			
Pro	Leu	Cys	Val	Glu	Leu	Asp	Glu	His	Asn	Leu	Pro	Arg	Phe	Pro	Glu
930				935				940							
Trp	Ile	Thr	Ile	Pro	Leu	Val	Cys	Ile	Tyr	Met	Leu	Ser	Thr	Asn	Ile
945				950				955				960			
Leu	Leu	Val	Asn	Leu	Leu	Val	Ala	Met	Phe	Gly	Tyr	Thr	Val	Gly	Thr
				965				970				975			
Val	Gln	Glu	Asn	Asn	Asp	Gln	Val	Trp	Lys	Phe	Gln	Arg	Tyr	Phe	Leu
				980				985				990			



<400>	779					
gattacgcaa	gctat ttagg	tgacactata	gaatwctcag	cttgcatcaa	gcttggtacc	60
gagctcggat	ccctagtaac	ggccgccagt	gtgctggaat	tgcgcccttgc	agccggggctc	120
agcatgagga	acagaaggaa	tgacactctg	gacagcacc	ggaccctgta	ctccagcgcg	180
tctcggagca	cagacttgtc	ttacagtga	agcgacttgg	tgaattttat	tcaagcaaat	240
tttaagaaac	gagaatgtgt	cttctttacc	aaagattcca	aggccacgga	gaatgtgtgc	300
aagtgtggct	atgcccagag	ccagcacatg	gaaggcacc	agatcaacca	aagtgagaaa	360
tggaactaca	agaaacacac	caagggaattt	cctaccgacg	cctttgggga	tattcagttt	420
gagacactgg	ggaagaaagg	gaagtatata	cgtctgtcct	gcgacacgga	cgcggaaatc	480
ctttacgagc	tgtcgaccca	gcactggcac	ctgaaaacac	ccaacctggt	catttctgtg	540
accggggggcg	ccaagaactt	cgccctgaag	ccgcgcacgc	gcaagatctt	cagccggctc	600
atctacatcg	cgcagtcctaa	aggtgcttgg	attctcacgg	gaggcaccca	ttatggcctg	660
atgaagtaca	tcgggggaggt	ggtgagagat	aacaccatca	gcaggagttc	agaggagaat	720
attgtggcca	ttggcatagc	agcttggggc	atggctctca	accgggacac	cctcatcagg	780
aattgcgatg	ctgaggggcta	tttttttagcc	cagtacctta	tggatgactt	cacaagagat	840
ccactgtata	tcttggacaa	caaccacaca	catttgcctg	tcgtggacaa	tggtgtgcat	900
ggacatccca	ctgtcgaagc	aaagctccgg	aatcagctag	agaagtatat	ctctgagcgc	960
actattcaag	attccaacta	tgggtggcaag	atccccattg	tgtgttttgc	ccaaggaggt	1020
ggaaaagaga	ctttgaaagc	catcaatacc	tccatcaaaa	ataaaattcc	ttgtgtggtg	1080
gtggaaggct	cgggccagat	cgctgatgtg	atcgctagcc	tgggtggaggt	ggaggatgcc	1140
ctgacatctt	ctgccgtcaa	ggagaagctg	gtgcgccttt	tacccgcac	ggtgtcccgg	1200
ctgcctgagg	aggagactga	gagttggatc	aaatggctca	aagaaattct	cgaatgttct	1260
cacctattaa	cagttattaa	aatggaagaa	gctggggatg	aaattgtgaq	caatgccatc	1320



```

tcctacgctc tatacaaagc cttcagcacc agtgagcaag acaaggataa ctggaatggg 1380
cagctgaagc ttctgctgga gtggaaccag ctggacttag ccaatgatga gattttcacc 1440
aatgaccgcc gatgggagtc tgctgacctt caagaagtca tgtttacggc tctcataaag 1500
gacagaccca agtttgtccg cctctttctg gagaatggct tgaacctacg gaagttttctc 1560
acccatgatg tcctcactga actcttctcc aaccacttca gcacgcttgt gtaccggaat 1620
ctgcagatcg ccaagaattc ctataatgat gccctcctca cgtttgtctg gaaactgggtt 1680
gcgaacttcc gaagaggctt ccggaaggaa gacagaaatg gccgggacga gatggacata 1740
gaactccacg acgtgtctcc tattactcgg caccctctgc aagctctctt catctgggcc 1800
attcttcaga ataagaagga actctccaaa gtcatttggt agcagaccag gggctgcact 1860
ctggcagccc tgggagccag caagcttctg aagactctgg ccaaagtga gaacgacatc 1920
aatgctgctg gggagtcgga ggagctggct aatgagtacg agaccggggc tgttgagctg 1980
ttcactgagt gttacagcag cgatgaagac ttggcagAAC agctgctggg ctattcctgt 2040
gaagcttggg gtggaagcaa ctgtctggag ctggcggtgg aggccacaga ccagcatttc 2100
atcgcccagc ctgggggtcca gaattttctt tctaagcaat ggtatggaga gattttccga 2160
gacaccaaga actggaagat tatcctgtgt ctgtttatta tacccttggg gggctgtggc 2220
tttgtatcat ttaggaagaa acctgtcgac aagcacaaga agctgctttg gtactatgtg 2280
gcgttcttca cctccccctt cgtggtcttc tcctggaatg tgggtcttcta catcgcttcc 2340
ctcctgctgt ttgctacgt gctgctcatg gatttccatt cgggtgccaca ccccccgag 2400
ctggctctgt actcgctggg ctttgtcttc ttctgtgatg aagtggagaca gtggtagcta 2460
aatgggggtg attattttac tgacctgtgg aatgtgatgg acacgctggg gcttttttac 2520
ttcatagcag gaattgtatt tcggctccac tcctctaata aaagctcttt gtattctgga 2580
cgagtcattt tctgtctgga ctacattatt ttactctaa gattgatcca catttttact 2640
gtaagcagaa acttaggacc caagattata atgctgcaga ggatgctgat cgatgtgttc 2700
ttcttctctg tcctctttgc ggwtggatg gtggcctttg gcgtggccag gcaagggatc 2760
cttaggcaga atgagcagcg ctggaggtgg atattccgtt cggctcatcta cgagccctac 2820
ctggccatgt tcggccaggt gccagtgac gtggatggta ccacgtatga ctttgccac 2880
tgcaccttca ctgggaatga gtccaagcca ctgtgtgtgg agctggatga gcacaacctg 2940
ccccggttcc ccgagtggat caccatcccc ctgggtgtgca tctacatgtt atccaccaac 3000
atcctgctgg tcaacctgct ggtcgccatg tttggctaca cgggtgggcac cgtccaggag 3060
aacaatgacc aggtctggaa gttccagagg tacttctctg tgcaggagta ctgcagccgc 3120
ctcaatatcc ccttcccctt catcgtcttc gcttacttct acatgggtgg gaagaagtgc 3180
ttcaagtgtt gctgcaagga gaaaaacatg gagtcttctg tctgctgttt caaaaatgaa 3240
gacaatgaga ctctggcatg ggaggggtgc atgaaggaaa actaccttgt caagatcaac 3300
acaaaagcca acgacacctc agaggaaatg aggcacgat ttagacaact ggatacaaag 3360
cttaatgatc tcaagggctc tctgaaagag attgctaata aaatcaaata aaactgtatg 3420
aactctaatz gagaaaaatc taattatagc aagatcatat taaggaatgc tgatgaacaa 3480
ttttgctatc gactactaaa tgagagattt tcagaccctt gggtagatgg tggatgattt 3540
taaatacccc tagtgtgctg agaccttgag aataaagtgt gaagggcgaa ttctgcagat 3600
atccatcaca ctggcgggccg ctogagcatg catctagag 3639

```

<210> 780

<211> 1095

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> (1)...(1095)

<223> Xaa = Any Amino Acid

<400> 780







Glu	Lys	Tyr	Ile	Ser	Glu	Arg	Thr	Ile	Gln	Asp	Ser	Asn	Tyr	Gly	Gly	
	275							280					285			
Lys	Ile	Pro	Ile	Val	Cys	Phe	Ala	Gln	Gly	Gly	Gly	Lys	Glu	Thr	Leu	
	290					295					300					
Lys	Ala	Ile	Asn	Thr	Ser	Ile	Lys	Asn	Lys	Ile	Pro	Cys	Val	Val	Val	
305					310					315					320	
Glu	Gly	Ser	Gly	Gln	Ile	Ala	Asp	Val	Ile	Ala	Ser	Leu	Val	Glu	Val	
				325					330					335		
Glu	Asp	Ala	Leu	Thr	Ser	Ser	Ala	Val	Lys	Glu	Lys	Leu	Val	Arg	Phe	
			340					345					350			
Leu	Pro	Arg	Thr	Val	Ser	Arg	Leu	Pro	Glu	Glu	Glu	Thr	Glu	Ser	Trp	
		355					360					365				
Ile	Lys	Trp	Leu	Lys	Glu	Ile	Leu	Glu	Cys	Ser	His	Leu	Leu	Thr	Val	
	370					375					380					
Ile	Lys	Met	Glu	Glu	Ala	Gly	Asp	Glu	Ile	Val	Ser	Asn	Ala	Ile	Ser	
385					390					395					400	
Tyr	Ala	Leu	Tyr	Lys	Ala	Phe	Ser	Thr	Ser	Glu	Gln	Asp	Lys	Asp	Asn	
				405					410					415		
Trp	Asn	Gly	Gln	Leu	Lys	Leu	Leu	Leu	Glu	Trp	Asn	Gln	Leu	Asp	Leu	
			420					425					430			
Ala	Asn	Asp	Glu	Ile	Phe	Thr	Asn	Asp	Arg	Arg	Trp	Glu	Ser	Ala	Asp	
		435					440					445				
Leu	Gln	Glu	Val	Met	Phe	Thr	Ala	Leu	Ile	Lys	Asp	Arg	Pro	Lys	Phe	
	450					455					460					
Val	Arg	Leu	Phe	Leu	Glu	Asn	Gly	Leu	Asn	Leu	Arg	Lys	Phe	Leu	Thr	
465					470					475					480	
His	Asp	Val	Leu	Thr	Glu	Leu	Phe	Ser	Asn	His	Phe	Ser	Thr	Leu	Val	
			485						490					495		
Tyr	Arg	Asn	Leu	Gln	Ile	Ala	Lys	Asn	Ser	Tyr	Asn	Asp	Ala	Leu	Leu	
		500						505					510			
Thr	Phe	Val	Trp	Lys	Leu	Val	Ala	Asn	Phe	Arg	Arg	Gly	Phe	Arg	Lys	
		515					520					525				
Glu	Asp	Arg	Asn	Gly	Arg	Asp	Glu	Met	Asp	Ile	Glu	Leu	His	Asp	Val	
	530					535					540					











```
<210> 786
<211> 45
<212> DNA
<213> Homo sapiens
```



<400> 786		
tacaccatcg	ggctgggcct gcacagtctt gaggccgacc aagag	45
<210> 787		
<211> 42		
<212> DNA		
<213> Homo sapiens		
<400> 787		
ttccagaact	cctacaccat cgggctgggc ctgcacagtc tt	42
<210> 788		
<211> 45		
<212> DNA		
<213> Homo sapiens		
<400> 788		
ctgtcagccg	cacactgttt ccagaactcc tacaccatcg ggctg	45
<210> 789		
<211> 45		
<212> DNA		
<213> Homo sapiens		
<400> 789		
catccgcagt	gggtgctgtc agccgcacac tgtttccaga actcc	45
<210> 790		
<211> 45		
<212> DNA		
<213> Homo sapiens		
<400> 790		
tcgggcgtcc	tggtgcatcc gcagtgggtg ctgtcagccg cacac	45
<210> 791		
<211> 45		
<212> DNA		
<213> Homo sapiens		
<400> 791		
aacgaattgt	tctgctcggg cgtcctggtg catccgcagt ggggtg	45
<210> 792		
<211> 45		
<212> DNA		
<213> Homo sapiens		
<400> 792		
gcactggtca	tggaaaacga attgttctgc tcgggcgtcc tggtg	45



<400> 793  
tcgcagccct ggcaggcggc actggtcatg gaaaacgaat tgttctgctc g 51

<400> 794  
atcagcattg cttcgcagtg ccctaccgcg gggaactctt gcctc 45

<400> 795  
tccgtgtccg agtctgacac catccggagc atcagcattg cttcg 45

<400> 796  
atcaagttgg acgaatccgt gtccgagtct gacaccatcc ggagc 45

<400> 797  
aacgacctca tgctcatcaa gttggacgaa tccgtgtccg agtct 45

<400> 798  
agacccttgc tcgctaacga cctcatgctc atcaagttgg acgaa 45

```
<210> 799
<211> 15
```







<400> 804

<210> 805

&lt;212&gt; PRT

<213> Homo sapiens

<400> 805

<210> 806

<211> 15

<212> PRT

<213> Homo sapiens

<400> 806

<210> 807

<211> 15

<212> PRT

<213> Homo sapiens

<400> 807

<210> 808

<211> 15

<212> PRT

<213> Homo sapiens

<400> 808

<210> 809

<211> 17

<212> PRT

<213> Homo sapiens



Ser Gln Pro Trp Gln Ala Ala Leu Val Met Glu Asn Glu Leu Phe Cys  
5 10 15

```
<210> 810
<211> 15
<212> PRT
<213> Homo sapiens
```

Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly Asn Ser Cys Leu  
5 10 15

```
<210> 811
<211> 15
<212> PRT
<213> Homo sapiens
```

Ser Val Ser Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser  
5 10 15

```
<210> 812
<211> 15
<212> PRT
<213> Homo sapiens
```

Ile Lys Leu Asp Glu Ser Val Ser Glu Ser Asp Thr Ile Arg Ser  
5 10 15

```
<210> 813
<211> 15
<212> PRT
<213> Homo sapiens
```

Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser  
5 10 15

```
<210> 814
<211> 15
```



<213> Homo sapiens

Arg Pro Leu Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu  
5 10 15